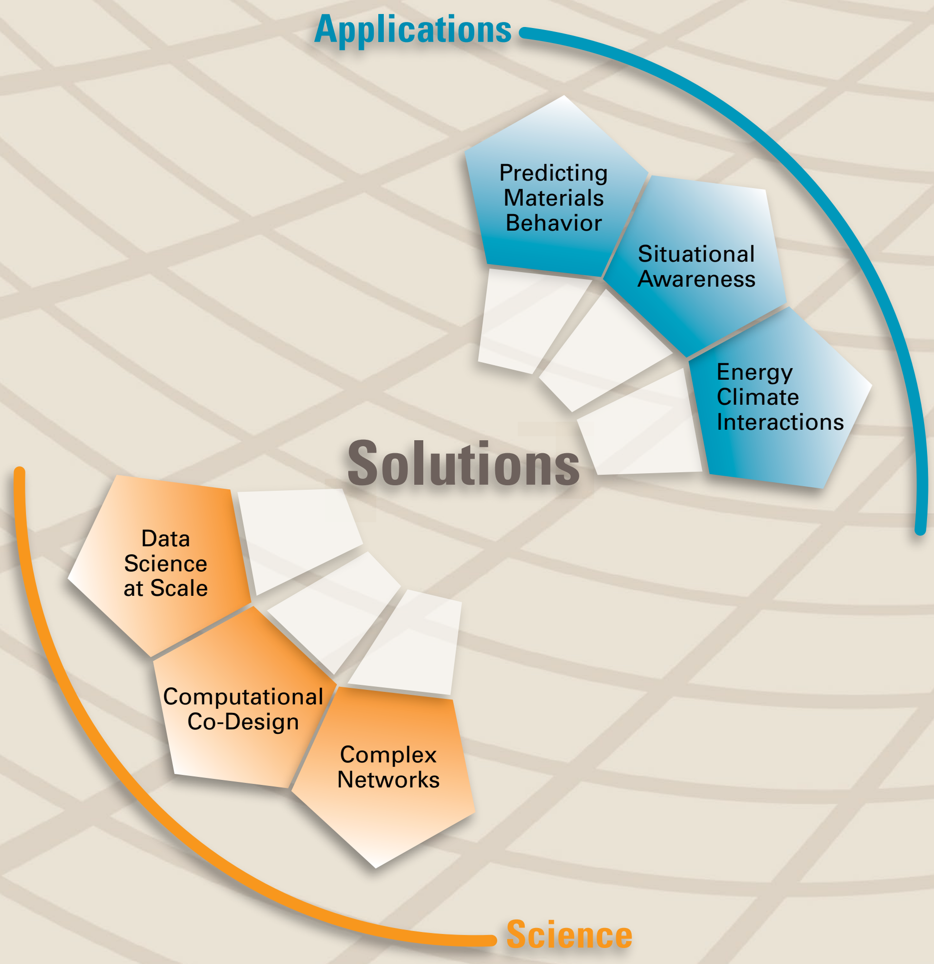
Science that Matters: Integrating Information, Science, and Technology for Prediction

Implementation Plan v 3.0
February 2014



Los Alamos
NATIONAL LABORATORY
EST. 1943

LA-UR-14-23474

Table of Contents

Executive Summary

LANL's capability pillar, Integrating Information Science and Technology for Prediction, addresses emerging challenges in national security, societal prosperity, and fundamental science. This pillar leverages advances in theory, algorithms, and the exponential growth of high-performance computing to accelerate the integrative and predictive capability of the scientific method. The Information Science and Technology (IS&T) pillar focuses on the integration of LANL assets for the understanding, quantified prediction, and design of complex systems.

The IST&T Pillar encompasses three capability areas: (1) Data Science at Scale, (2) Computational Codesign, and (3) Complex Networks.

(1) The Data Science at Scale capability seeks to exploit the extremely large datasets and extremely high-rate data-streams from sensors, embedded computing and traditional high-performance computing. The Data Science at Scale capability provides a set of tools capable of making quantifiably accurate predictions for complex problems with efficient use/collection of data and computing resources. Analysis of these datasets is widely recognized as a new frontier at the interface of information science, mathematics, computer science, and computer engineering.

(2) Computational Codesign leverages the breadth of LANL capabilities as expressed through integrated multi-physics computational software, providing numerical solution and insights into a wide variety of grand challenge problems. Given our strengths in the creation of multi-physics computational physics software and analysis tools, LANL has created a new field of computational co-design intended to address the significant challenges due to disruptive change in computational infrastructure. Computational co-design addresses this challenge through the collaborative development of hardware, system software, algorithms, and applications with each informing the other so as to enable the best performance given a set of constraints in terms of price, performance, power, and resilience.

(3) Complex Networks provide a natural framework for describing complex systems in terms of interdependent subsystems and components. Since complex systems and their network representations arise in many scientific and national security mission areas at LANL, it is imperative that we build the scientific foundation for work in this area. Cyber systems, national infrastructure, biological systems, real and artificial neural systems, classical and quantum communication systems, systems of networked sensors, social networks, terrorist networks, and energy systems such as the smart grid, all rely on a scientific foundation of complex networks.

Challenges for FY14 and Future

Over the past year, we have identified demand expected growth trajectories for a few key areas and will focus our efforts and investments in those areas in FY14.

Data Science at Scale

The urgent need for a capability in large-scale data fusion has been identified across a wide swath of programs including biosurveillance, nuclear nonproliferation and other applications relevant to LANL's Science of Signatures Pillar. To meet this growing need, in FY14 we will pursue pilot projects in biosurveillance and in nuclear nonproliferation. Key to effective decision support resulting from such fusion is the ability to assess quantitatively the uncertainty in the output of fused data.

Computational Codesign

LANL's first instantiation of Codesign (i.e., Codesign 1.0) involved the optimal design of algorithms/software/hardware systems. In FY14, building on our successes in Office of Science and LDRD projects we will integrate our newly developed capabilities directly into code projects within the ASC program. We will also emphasize investment in the development of predictive multiscale and multiphysics methods which live at the convergence large-scale data analysis and exascale computing.

Complex Networks

For Complex Networks we will leverage recent successes of our Smart Grid efforts and expand our scope to a broader set of cyberphysical systems.

Capability Enablers

Partnerships

In late FY13 the new set of PSAAP Centers was announced and we will explore tight alliances with a number of the participating universities including Stanford and Texas A&M.

Key to success in IS&T, as in other ventures, is the ability to partner with external leaders in the field. These can be at both the individual investigator level or at a broader, institutional level, and include academic, industrial, and other governmental entities. The Information Science and Technology Institute (ISTI) is a particular conduit of such relationships in support of the IS&T pillar. ISTI will continue to support partnerships with:

- UC Davis's Institute for Next Generation Visualization and Analysis
- UCSC Institute for Scalable Scientific Data Management
- Carnegie Mellon University
- Missouri S&T
- MIT
- NYU Center for Urban Science and Progress (CUSP)

In addition ISTI will establish formal partnerships with Johns Hopkins, Iowa State, Texas A&M and Stanford Universities.

LANL will continue to support the Information Science and Technology Institute and the Center for Nonlinear Studies. Both of these organizations have been integral to the implementation of the IS&T Pillar.

Workshops and Summer Schools

Workshops and conferences provide LANL branding, demonstrate LANL leadership, and help to formulate a path forward, while cutting across capabilities and program applications. Despite constraints on travel we will continue to support these events for scientific exchange. Our schedule for FY14 includes:

- Materials Informatics-Feb 2014
- Conference on Data Analysis (CODA)- 2014
- Quantitative Biology -August 2014
- Systemic Risk- Fall 2015

We will continue to support our highly successful Summer Schools:

- Codesign
- Cluster
- Big Data

With the successes of the past year and with the plans for FY14 outlined in this document, we believe that the IS&T Pillar has an exciting future addressing some of the nations most pressing challenges. In Spring 2014, LANL will hold the third IKS Capability Review. This review (one of 3 representing the IS&T pillar) will focus on Data Science and Complex Networks and Applications of the capabilities in those areas.

Introduction

Los Alamos National Laboratory's IS&T capability builds upon our decades-long expertise and rich tradition in integrating high performance computing, theory, statistics, learning, modeling, simulation, and the instrumentation and diagnostics of scientific experiments. Our national security and economic competitiveness require that the United States remain at the forefront of IS&T. The LANL capability pillar, Integrating Information, Science, and Technology for Prediction will directly support DOE's mission to deliver capabilities to scientists nationwide that enable them to extend the frontiers of science, and will provide the ongoing support for the US nuclear deterrent.

The IS&T Strategic Plan was released in early 2011 and Versions 1.0 and 2.0 of the Implementation plan were released in 2011 and 2012 respectively. In FY13 the IS&T pillar has been further refined and has evolved into a more coherent structure. Over the past year there have been a number of structural/organization changes to the IS&T Pillar. Several new themes were established. These themes include Programming Models and Machine Learning (replacing Data driven modeling). The Scale-bridging and Extreme N-body Solvers themes were eliminated and their scope moved to the Computational Physics and Computational Mathematics themes respectively. New leads were appointed for several themes. Themes with targeted applications were moved from the three aforementioned capability areas to a separate Applications Area. These domain-specific themes are AstroInformatics, BioInformatics, Materials Informatics, Climate Informatics, Smart Grid, HPC Operational Informatics, and Structural Health Informatics. As the other science pillars (Materials for the Future and Science of Signatures) have evolved, the IS&T Pillar has evolved along with them ensuring a strong alignment into the future.

This document is a detailed implementation plan explaining the path to achieving the goals outlined above and in the Integrating IS&T for Prediction Strategic Plan (LA-UR 11-00115). In addition to a new section on applied informatics, as in the Strategic Plan, we:

- List and review the strategic goals of the application areas: Predicting Materials Behavior, Situational Awareness, and Energy-Climate Impact and Energy Infrastructure
- Define the scope and objectives of the three capability areas: Data Science at Scale, Computational Co-design, and Complex Networks and the Applications Area. Each of these areas is further divided into focused themes described in this document

This implementation plan outlines the steps and details requiring resources for executing the IS&T strategy. This plan is an evolving document and will be modified over the course of time to meet Laboratory and national needs.

Program Areas

Predicting Materials Performance

The overarching IS&T grand challenge in materials modeling, is to develop a set of tools to predict and eventually to tailor, the response of a material. As outlined in the Integrating IS&T for Prediction Strategic Plan, the goals for predicting materials performance are:

- Develop capability to achieve optimal use of data and design of experiments for the enhanced quantitative prediction of materials performance
- Establish the technical underpinnings of data-driven analysis and compare to results from simulations (in, for example, reconstructing microstructure from diffraction data) and quantified predictive capability for materials performance
- Enable the vision of co-design through the effective integration of fundamental theory, multiscale computation, and experiment
- Exploit the emergence of exascale-class computing by pursuing co-design strategies that integrate physics models, programming platforms, and hardware for bridging scales using adaptive physics refinement
- Pursue applications of complex networks science to the description and model solutions for the interaction of disordered structures within materials

These efforts are entirely consistent with LANL's, Experimental Science Focused on Materials for the Future Pillar and strategy, and the two pillars are mutually supportive.

Situational Awareness

Situational Awareness is the perception of elements within a potentially large domain of time and space, the comprehension of their meaning, and the projection of their status into the near future. It concerns perception of the environment critical to decision makers in complex, dynamic areas such as nonproliferation, counterterrorism, military command and control, space, cyber infrastructure, and emergency response.

The fundamental challenge for information science is to expose, extract, and predict phenomena of interest from an increasingly diverse set of sensors and information sources and from an increasingly cluttered and unpredictable background. To address this challenge we will require computational solutions that draw from a diverse set of capabilities that include sensor modeling, modality-specific data modeling, analytics modeling, knowledge modeling, and large-scale data management, as well as visualization, human computer interaction, and decision support sciences.

The goals for situational awareness include the following:

- Develop and execute a beginning-to-end pilot project for a specific global security application to demonstrate LANL capabilities in optimal inference/situational awareness
- Provide scientific underpinnings for cybersecurity research via interdisciplinary collaborations and programmatic opportunities

- Grow the cybersecurity component of LANL's Quantum Information Systems (QIS) capability within the National Nuclear Security Administration (NNSA) and other government organizations
- Maintain and further develop LANL's world-class expertise in specific technical components relevant to situational awareness, such as sensor, algorithm, and application niches
- Provide the scientific foundation for unifying situational awareness capacities across the Laboratory
- Focus on coupling the ability to collect large quantities of data with emerging IS&T tools in order to turn this data into information and then actionable knowledge; form cross-disciplinary and cross-laboratory teams to achieve this coupling

These efforts are entirely consistent with LANL's Science of Signatures Pillar and strategy, and the two pillars are mutually supportive.

Energy-Climate Impacts and Energy Infrastructure

Society's energy use and its impacts on our global and regional climate pose scientific, technological, and societal challenges. The IS&T aspects of these challenges will involve efforts in the following areas:

- Quantification of greenhouse gas (GHG) emissions and offsets for policy makers and scientists
- Uncertainty quantification (UQ) for global and regional climate prediction
- Infrastructure modeling, design, and optimization

These IS&T efforts are entirely consistent with LANL's Science of Signatures Pillar and the Experimental Science Focused on Materials for the Future Pillar and all three pillars are supportive of one another.

The goals for energy climate impacts and energy infrastructure are:

- Develop scientific underpinnings for IS&T aspects of smart grid
- Work with climate scientists to develop advanced UQ capability for climate prediction
- Foster development of cross-disciplinary data-fusion capability, including climate-energy-geospatial systems
- Participate in the emerging discipline of energy impacts science through the Energy Climate Impacts program development effort
- Develop new methods for quantifying the uncertainties in GHG for the program development effort tasked with developing a new Greenhouse Gas Information System (GHGIS)

Capabilities and Themes

Capability: Data Science at Scale

The goals for data science at scale are:

- Position LANL to compete for external program opportunities involving massive datasets and streaming data by:
    - Fostering the scientific underpinnings of data-intensive computing, including streaming data
    - Launching a data science at scale laboratory for hardware/software/algorithmic experimentation with massive datasets
    - Building scalable data-processing capabilities in statistics, UQ, and machine learning
    - Collaborating with strategic external partners in the area of data science
- Position LANL to compete for external program opportunities requiring D2D by:
    - Establishing practical yet rigorous detection science as part of a targeted global security program
    - Developing a nationally recognized program and becoming the lead national laboratory in inference and machine learning and UQ
    - Building integrated efforts that jointly solve problems in the area of D2D (for example,visualization/machine learning, statistics, and data infrastructure

Theme: Data Management and Curation

Massive amounts of research data are being created (on a continuum of size from multiple small datasets to huge streaming data), but scientists, data managers, and librarians are in their formative stages when it comes to acquiring critical 21st century skills. Some of the challenges are how to best manage this data so it can be exploited, validated, re-used, and repurposed. Developing standards and best practices, and making available infrastructure that can be leveraged during the entire data lifecycle, is critical to achieve these goals. Recording appropriate provenance information is an essential ingredient of any possible solution.

Creating the appropriate “eco-system” to support robust management of research data is not a trivial endeavor. Many of the challenges involved require broad collaboration, both at the institutional level (across institutional resources, including IT, HPC, CIO, library, and researchers), at the level of scientific disciplines, and at the international level (cf. the EC EUDAT project and the global Research Data Alliance). Understanding that federal agencies - including NSF, NIH, and DOE – have started requiring certain guarantees regarding the management and preservation of digital data as a contingency for funding, it is essential to become active in a concrete manner in this realm.

In March 2012, the US Office of Science and Technology Policy issued a “Big Data Research and Development Initiative.” Among the goals is to Advance state-of-the-art core technologies needed to collect, store, preserve, manage, analyze, and share huge quantities of data. Many projects are underway at LANL, in the DOE, across the nation, and across the globe to tackle certain aspects of the problem. These include the development of new computational paradigms and software for processing and analyzing large volumes of data and to visualize them in comprehensible ways, and approaches to automatically capture and store provenance data in workflow system environments.

Other areas, however, are underserved, most notably the tasks around developing a vision and strategy for building the needed infrastructure to manage, curate, and serve data for re-use and future exploitation.

Working forward, we must urgently develop competencies, policy, and shared infrastructure for data-intensive science at LANL. This involves building 4 core areas:

- Data literacy (education)
- Development of data management best practices regarding:
  - Metadata about data (identification, aggregation, versioning, annotation, authorization, embargo, etc.)
  - Metadata about the processes applied to data (archiving of workflow descriptions, archiving of descriptions of services involved in workflows, archiving of code involved in workflows, etc.)
  - Data formats
- Policy and practice to position LANL so it can be a successful, competitive player in the realm of data-intensive research
- Shared infrastructure for data staging, exploitation, and archiving

This theme area must intersect with other IS&T projects/themes to determine where best to embed data management research and practical questions to leverage library/information science capabilities.

FY14 Action Items

- Launch a collaborative effort within LANL, involving all major constituents, aimed at defining, designing, and prototyping shared research data infrastructure for open science with an eye on both exploitation and preservation of data
- Engage in the efforts of the Research Data Alliance (https://rd-alliance.org/), a global effort that works towards establishing interoperable infrastructure for data sharing

Existing Efforts

In FY12, two LANL staff became official members of the NSF DataONE working groups which engages LANL with a large, multi-institutional effort developing practices and policy around data management.

- The Research Library conducted a survey in early FY12 targeting authors and NSF grantees to learn status of data management efforts. The Library has the results.
- The Research Library started activities pertaining to Data Management Plans.
- The Research Library organized a meeting and seminar around data management challenges with representatives from the US DataONE and the EC EUDAT project.
- The Research Library's Herbert Van de Sompel acts as external advisor to the EC EUDAT and iCORDI projects that operate in the realm of interoperable data infrastructures.
- The Research Library is engaging LANL efforts regarding research data management.

Required Resources

- In order to launch a LANL effort to explore shared research data infrastructure for open science, a working group with representation from major LANL constituents should be set up under the auspices of PADSTE. Members of the working group need to commit a sizeable portion of their time to this activity and, for explorative and prototyping activities, will need support from skilled IT staff.
- In order to engage in the Research Data Alliance, a LANL representative needs to be able to dedicate a sizeable portion of his/her time to activities in select Working Groups (see https://rd-alliance.org/working-and-interest-groups.html) and to attend the Plenary Meetings, which, so far, have been organized twice a year at an international location.

Theme: Data to Decisions (D2D)

The goal of Data to Decisions (D2D) is to make information useful by converting raw data into actionable knowledge. Without a strong D2D capability, data collected from sensors or stored in databases, amounts to disconnected dots, and it is both difficult and time-consuming for decision-makers to gather the information they need in order to connect the dots, see patterns, and solve problems. With a strong D2D capability, organizations and individuals can access and use information efficiently and effectively.

D2D is thus the goal of the information flow that links information sources with information users. This information flow can be conceptualized as a pyramid (Fig. 2) that combines other capabilities, such as information collection, extraction, integration, representation, and dissemination. Research conducted at LANL addresses all of these problems and includes data collection from space, air, and ground, information extraction from image and text, knowledge integration in mission-critical domains, and representation of actionable knowledge in a form optimized for users. A key challenge is to combine these capabilities in such a way that information from diverse and distributed sources passes through the pipeline(s) rapidly, with no loss in accuracy, and the users receiving the information they need in a form they can use. The D2D Implementation Plan outlined below therefore focuses on connecting the dots within LANL in order to more fully develop our D2D capabilities.

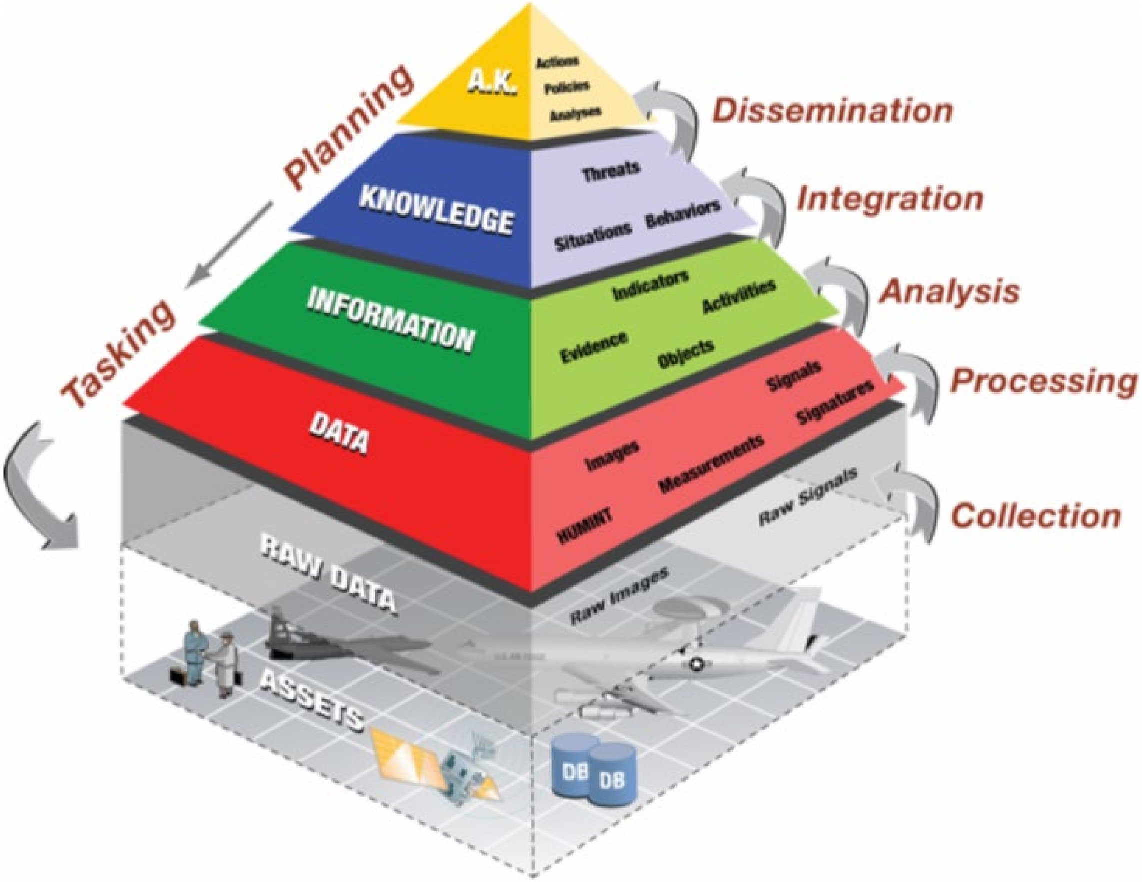


Figure 2. Information flow pyramid

Milestones and Goals for FY14

- Apply the D2D capabilities identified in the first phase of the Strategic Outcomes Office Grand Challenge in "Loose Nukes" to the second phase of the exercise
- Work with the SOO principals and crosscuts to develop another grand challenge involving D2D, persistent surveillance, and space defense
- Identify the D2D capabilities that apply to the new grand challenge
- Work with the PIs to apply their capabilities to the grand challenge
- Work with Program Managers to identify sources of potential funding
- Work with PIs to write proposals to obtain funding

Milestones and Goals for FY15

- Work with the SOO principals and crosscuts to develop another grand challenge
- Identify the D2D capabilities that apply to the new grand challenge
- Work with the PIs to apply their capabilities to the grand challenge
- Work with Program Managers to identify sources of potential funding
- Work with PIs to write proposals to obtain funding

Milestones and Goals for FY16 and Beyond

- Work with the SOO principals and crosscuts to develop another grand challenge
- Identify the D2D capabilities that apply to the new grand challenge
- Work with the PIs to apply their capabilities to the grand challenge
- Work with Program Managers to identify sources of potential funding
- Work with PIs to write proposals to obtain funding

Required Resources

- FY 14 – Funding for the Grand Challenges from the SOO, PADGS, PADSTE, and LDRD
- FY15 – Funding for the Grand Challenges from the SOO, PADGS, PADSTE, and LDRD
- FY16 and beyond – Funding from sponsors

External Partners

- Sandia National Laboratory (SNL), potential
- Lawrence Livermore National Laboratory (LLNL)

Existing Efforts

- The SOO Grand Challenge in "Loose Nukes"
- Discussions with the SOO about another grand challenge involving D2D, persistent surveillance, and space defense

Theme: Image and Signal Processing

LANL's interest in image analysis and signal processing spans a wide range of applications, from microscopy for nuclear forensics to remote sensing for nonproliferation. Areas of research include video surveillance, time-domain and frequency-domain, radio-frequency (RF) signal analysis, multispectral and hyperspectral imagery analysis, target detection, anomaly detection, change detection, image co-registration, 3D reconstruction, compressive sensing, computational imaging, sparse modeling, single-photon imaging, hierarchical modeling, and image segmentation.

Image analysis and signal processing are relatively mature fields in IS&T. There is considerable practical expertise at LANL in using these image analysis and signal processing tools, and that expertise is distributed across the Laboratory. However, advancing the Laboratory's missions will require more than the standard toolkits. Data sets are orders of magnitude larger and increasingly heterogeneous, models are more sophisticated, systems are more complicated, and the world is just not as Gaussian as it once seemed to be. Recent advances in machine learning, sparse modeling, and biomimetic engineering have provided new opportunities for dramatically enhancing the tools we use for image analysis and signal processing.

Whether it is monochromatic, traditional red-green-blue, multispectral, or hyperspectral, image datasets present the analyst with a rich spatio-spectral structure. For multiple images, and particularly for video, the temporal aspect adds a further dimension of richness. Interpretation of imagery is something for which humans are naturally adept and computers are notoriously inept.

As one writer put it: "Computer vision? I'll believe it when it sees me." But even small steps can be valuable, as the Genie project demonstrated. The combination of machine learning and image analysis provided the basic underpinning of the Genie project, this work led to award-winning (R&D100) software that was commercially licensed both for remote sensing and for medical imaging applications.

In general, a good source of opportunities for major advances in a mature field is by synergistic combination with new and emerging fields. Sparse modeling is such a field, and it comprises one component of the ongoing LDRD/DR project on "Hierarchical sparse models for robust analysis of video data." An emphasis on geospatial applications is another force that is pushing the technology of image analysis.

Milestones FY14

- Publish the results of funded work in the conference and peer-reviewed literature
- Invite external Image and Signal Processing scientists and engineers to the IS&T Speaker Series, and cultivate interactions and collaborations with them
- Identify and pursue funding for image and signal processing research

Stretch Goals

- Understand what is in an image, and what is going on in a video
- Provide tools to more accurately extract specific information from scientific and programmatic imagery
- Effectively fuse image and signal data from multiple sensors
- Develop ways to effectively incorporate domain expertise into automated image and signal processing
- Improve collaboration between image/signal acquisition (i.e., sensor design) and image/signal processing
- Scale up processing tools to permit the analysis of increasingly larger datasets
- Spin-off existing capabilities by making them more tangible, accessible, and usable to our customers

Existing Efforts

- Geospatial semantic analysis
- Anomaly, target, and change detection
- Image segmentation and characterization for forensic material science
- Sparse modeling of time series signals with both fixed and adaptive bases
- LDRD/DR project "Hierarchical sparse models for robust analysis of video data"
- LDRD/DR project "Empowering the Expert: Machine Learning with User Intelligence" includes adaptive image segmentation as one application

External Partners

- Goodrich, commercial CRADA for change detection
- Observera, provides a stand-alone Genie Pro
- Leica Biosystems (formerly Aperio), licenses Genie Pro for medical imaging
- Invertix
- UCOP (University of California Office of the President)
- Other National Laboratories, such as LLNL

Theme: Machine Learning

Machine learning lies at the intersection of computer science and statistics, and provides an important set of tools for a wide range of applications including, signal processing, bio-informatics, materials informatics, cyber security and knowledge management. This summary is divided into two parts. The first part describes existing capabilities that were developed through past projects. Many of these capabilities are now standard tools in a number of applications at LANL and in the larger information science community. The second part describes new capabilities that are in development (currently funded), or are potential targets for development in the near future.

Core Machine Learning Capabilities

Classification: Classification is one of the most important, and fundamental, problems in machine learning. It focuses on finding functions that predict binary labels from data. The problem is important because it has a large number of applications. The problem is fundamental because many (more complex) problems can be reduced to binary classification. Understanding (and quantifying) the performance of different methods therefore provides a firm theoretical basis for machine learning in general. Over the last decade LANL has been at the forefront of these theoretical advances.

Non-Traditional Problem Reductions: LANL has also been at the forefront of reducing different applications into classification problems. Detecting rare events is critical to many applications and LANL has helped develop formal mappings that reduce this problem to the classification problem. This mapping has been developed at LANL for applications in cyber-security, materials science as well as content based search of images. LANL has also pioneered the reduction of classification methods to change detection. In this case, the function predicts labels from two different datasets, often taken at two different points in time. Other reduction examples include Min-Max Robust Classification, Neyman-Pearson Detection, as well as Hemi- and Semi-Supervised Learning paradigms. There are a large number applications that could potentially benefit from these reductions including fault detection, satellite state-of-health monitoring and cyber security.

Feature Extraction / Selection: To successfully apply machine learning in practical applications requires a large application-specific effort to choose, and often hand craft, the data that is input to the machine learning system. LANL has been at the forefront of feature extraction / feature selection methods that aim to reduce this up-front cost, and produce more general purpose algorithms. For image and signal processing, LANL has advanced the state-of-the-art in compressed sensing, and developed methods that are general purpose and yet achieve state-of-the-art performance.

Emerging machine learning capabilities

Machine Learning meets Graphical Models: A significant new thrust in machine learning is to extend prediction methods from labels to more complex outputs such as sequences, parse trees, and more generally, dynamic typed graphs. This thrust has many names (and specializations) including structured output prediction, statistical relational learning and conditional random fields. This thrust has required machine learning to unify with other areas of information science including Bayesian Networks, Graphical Models, Inductive Logic Programming and Reinforcement Learning. This unification promises many new applications (some of which are discussed below) and presents many new challenges for capability development.

LANL is well positioned to advance this new thrust since it has core capabilities in machine learning (discussed above) and also existing capabilities in the theory and application of Graphical Models. Theoretically, LANL has been at the forefront of Graphical Models due to its close connection with statistical physics and Markov Chain Monte Carlo optimization. In this context Graphical Models have been developed in depth for (smart) power grid applications. More practically, LANL has also been a leader in applying Graphical Models to Global Security applications with an emphasis on decision support systems. The unification of machine learning and Graphical Models can benefit applications in two main ways: 1) it provides a mechanism to incorporate domain knowledge into machine learning methods, and 2) it enables Graphical Models (prior knowledge) to adapt to data.

Knowledge rich models: Incorporating domain knowledge into machine learning methods could be a significant general purpose capability for LANL. This is because in addition to core machine learning capabilities, LANL has significant expertise in a variety of domains that could benefit from machine learning. These include basic science fields such as material informatics, bio informatics, complex system modeling and computational physics. It also includes more applied fields related to weapons, non-proliferation, arms control, and space situational awareness. A challenge for LANL (and machine learning in general) is to develop methods that can efficiently formalize and exploit this domain expertise within machine learning methods. Feature Extraction and Structured Output Prediction are important (and complementary) components that can address this challenge. Successful demonstrations of such capabilities would highlight LANL's ability to exploit state-of-the-art machine learning for new and emerging applications in science and defense. It would also support LANL's continued role at the forefront of machine learning.

Interactive machine learning: This emerging capability focuses on the second benefit of structured output prediction: it enables machine learning methods to adapt to domain experts in the deployed environment. LANL has been at the forefront of interactive machine learning methods based on core capabilities for image analysis. By building on more recent advances in structured output prediction, interactive machine learning can benefit a much larger number of knowledge rich applications at LANL. Interactive machine learning and knowledge rich models provide complementary capabilities for many applications: knowledge rich models capture domain expertise that can be explicitly formalized, and interactive machine learning captures implicit domain knowledge that is captured through observation.

Decision/game theory: A promising approach to many of the difficult search/optimization problems encountered in many applications is *look-ahead*, where the benefit/risk of decisions (classifications) are estimated prior to making the decision. LANL, and collaborators, are at the forefront of this approach, and there are a variety of applications at LANL that could benefit, particularly if this approach can be combined with knowledge rich systems to provide decision support, or guide experimental design. LANL has also been actively developing methods for the more difficult multi-agent decision problem (i.e. game theory) and these can provide more accurate decision models in many practical applications e.g. cyber-security.

Machine Learning at Scale: As data volumes increase, and machine learning methods grow in complexity, there is an increasing need to leverage HPC. As HPC advances, there is a growing need to leverage decentralized and robust computational components. This has led to growing interest in methods that can exploit the complementary strengths of machine learning and HPC. LANL has developed considerable capability in a number of specific applications through HPC focused co-design efforts. These could help LANL build a more general purpose machine learning/ HPC capability that would benefit a large number of applications, and also help support LANL's continued role at the forefront of machine learning.

Milestones FY14

- Develop theory and algorithms for graphical models, belief propagation and inference
- Develop competence in structured output prediction, and knowledge based machine learning
- Publish the results of funded work in the conference and peer-reviewed literature
- Identify priority application areas for machine learning at LANL, such as nuclear material forensics, cyber-security, nonproliferation relevant signal processing as well as space situational awareness

Milestones FY15

- Establish capability in interactive machine learning, demonstrate technical leadership in this area, and engage research community and potential customers
- Establish connections between machine learning and other LANL capabilities such as optimization, network science, simulation and modeling as well as HPC
- Develop proposals, partnerships and projects for machine learning at LANL based on technical capabilities and priority applications

FY16 and Beyond

- Unification of machine learning, image and signal processing, information exploitation, digital data sciences and other data analytic / data mining capabilities under a single, lab–wide data modeling / exploitation capability. Internal to LANL this could be used to build more effective collaborations with applications, domain experts, visualization and HPC capabilities. External to LANL, this could provide a clearer (but perhaps over-simplified) representation of data analysis capabilities.
- Historically, machine learning at LANL has benefited greatly from a strong collaboration between theory and practice. Over the last ten years, machine learning at LANL has moved from a strong theoretical capability towards a more applied capability. A strong machine learning capability will require reinvigorating the theoretical development over a broader suite of application areas.

Ongoing LDRD Efforts

- LDRD-DR FY14-FY15 Material informatics
- LDRD-DR FY13-FY15 Machine Learning with User Intelligence
- LDRD-ER FY13-FY15 Planar Graphical Models
- LDRD-DR FY10-FY13 Cyber-security
- LDRD-ER FY10-FY13 Rare Category Detection

Resources Required

- Continued support from the IS&T Speaker Series, and cultivate interactions and collaborations with external scientists and engineers
- Strategic hires, particularly in the form of postdoctoral conversions, to help LANL maintain a current capability
- Continued IS&T institute support for collaborations with relevant academic and research centers and commercial entities in machine learning e.g. Lab fee and other UC programs
- Support for exploratory activities, particularly for building new internal collaborations between machine learning researchers and application experts and HPC staff

External Partners

- Academic: MIT, UC Irvine, UCSC, CMU, UT, UNM
- Commercial/Research centers: MudCube, Aperio, Jain Labs, Kitware Inc.

Theme: Programming Models and Architectures

Programming models are typically focused on achieving increased developer productivity, improved performance, and portability to other system designs. The rapidly changing nature of processor architectures and the complexity of designing complex computing platforms provide significant challenges for each of these goals. Several other factors are likely to impact the design of future programming models. From the aspect of computational science, the representation and management of increasing levels of parallelism, concurrency and memory hierarchies, combined with the ability to maintain a progressive level of interoperability with today's applications are of significant concern. While large-scale data science will also be impacted by changing processor architectures it is important to realize that the programming models, as well as underlying system architecture and supporting infrastructure, for data-centric computation may significantly diverge from those traditionally used in high-performance computing. This includes database-like models (such as query languages) and instances such as the now common map-reduce model made popular by Google and languages like ECL from Lexis-Nexis. Several significant challenges in this space arise when considering the intersection of computational, experimental and data sciences with the design of underlying system architecture, network infrastructure, and supporting software technologies, this is especially so when considering both productivity and performance with the explicit goal of increasing the rate of scientific discovery.

Milestones and Goals FY14

- Identify opportunities
- Hold a series of meetings that bring together program managers and research staff with the two-fold aim to educate the staff about the programmatic needs and opportunities in the space of architectures, infrastructure, and programming approaches
- Submit associated LDRD DRs and ERs
- Evaluate and develop strategic growth areas; strengthen collaborators with industry and academia

Milestones and Goals FY15

- Continued identification of opportunities and growth areas from programmatic missions
- Submission of associated LDRD DRs and ERs as well as other programmatic sources
- Evaluate and develop strategic growth areas; leverage established collaborators with industry and academia to further strengthen capabilities

Required Resources

- FY14 - Funding from LDRD Reserve, Program Development, Programmatic, and/or general and administrative (G&A)
- FY14 – Institutional and/or program funds to build data- and/or cloud-centric computing testbeds and necessary updates to LANL-wide infrastructure
- FY15 - Funding from LDRD DR or ER, Programmatic, and/or G&A
- FY15 – Institutional and/or program funds to build architecture-centric computing testbeds and necessary updates to LANL-wide infrastructure
- FY16 and beyond – Funding from LDRD DR or ER, Programmatic, and/or general and administrative (G&A)
- FY16 and beyond – Institutional and/or program funds to build data, and/or cloud-centric computing testbeds and necessary updates to LANL-wide infrastructure

External Partners

- Johns Hopkins University
- Stanford University
- IBM
- NVIDIA
- ARM

Theme: Statistics and Uncertainty Quantification

The Statistics and UQ theme includes an established base of statistical knowledge and experience coupled with a broad interdisciplinary body of expertise in uncertainty quantification. While the field of statistics inherently involves understanding and quantifying uncertainty, UQ typically refers to methods for understanding uncertainty present in modeling problems.

The LANL theme in Statistics and UQ addresses a broad set of issues important to multiple Laboratory missions including national security and global security, as well as emerging scientific issues in energy, materials, climate, and environment. Historically, a significant portion of the work in statistics and UQ has been focused on supporting the nuclear weapons program. More recently, UQ initiatives supporting broader security missions such as nonproliferation, homeland security, and biosecurity have emerged. The DOE's increased focus on energy security and development of new technologies for cleaner and more efficient energy has led to participation in innovative simulation programs such as the Consortium for Advanced Simulation of Light water reactors (CASL) and the multi-lab Carbon Capture Simulation Initiative which focuses on the development of cleaner technologies for coal-fired power plants. LDRD-funded work on cosmology to understand the origins of the universe and DOE-funded work for the ASCEM environmental subsurface modeling program are also benefitting from these approaches. Statistics and UQ play an important role in LANL's Materials mission, including work for both weapons physics and weapons engineering, and combining theory with data from both physical experiments and computer simulations. Recent initiatives have been established to explore the use of UQ in the area of nuclear forensics. As LANL's mission changes over time, the statistics and UQ capability is structured to respond to dynamic needs of importance to the nation.

Statistics

Statistics includes the development and application of a broad range of mathematically rigorous methods for exploring, summarizing, and analyzing data, and for drawing inferences based on data. The current LANL portfolio of statistical work includes a major focus on nuclear weapons, with significant efforts in global security, energy, and science. Within this mission space, state-of-the-art statistics is incorporated into collaborative efforts that address challenges in the planning and collection of data, summary and analysis, modeling and prediction, and statistical inference to support informed decision-making.

Milestones FY14

- Identify existing capabilities and potential areas for growth and innovation
- Develop materials describing technical capabilities, current programs, and expertise
- Identify key elements of computational infrastructure required to support statistical research and innovation

Milestones FY15

- Develop a strategic plan including goals, directions of emphasis, resource requirements, program development, staff development, and long-range planning for recruitment and retention
- Develop a resource plan to support state-of-the-art statistical computation and analysis of large data
- Examine hiring strategies to encourage recruitment of top talent needed in an interdisciplinary mission-driven research laboratory environment
- Strengthen and extend network of collaborators

Milestones FY16 and Beyond

- Become a leader in computational statistics and analysis of large data

Stretch Goal

- Provide the premier interdisciplinary environment for statistical research and innovation

Required Resources

- Stable portfolio of high-quality funding for existing and future staff
- Computational infrastructure
- Technical leadership
- Involvement of scientific staff
- Support from higher levels of management in pursuing high-quality sustainable funding and implementation of staffing plan

External Partners

- Iowa State University
- Other academic collaborators

Existing Efforts

- Numerous academic collaborations
- Participation in external professional conferences and activities
- Development of programmatic opportunities to support high-quality technical work for existing staff
- Second Conference on Data Analysis exploring data-focused research across the Department of Energy (February 2014)
- Exploring sponsorship of national research conference

Uncertainty Quantification

UQ involves the structured application of a suite of tools to identify and quantify sources of uncertainty in the modeling process. The UQ process provides approaches for synthesizing output from models (often involving large-scale computation) with experimental/observational data for the purpose of making predictions and understanding underlying phenomena. In addition to the actual predictions, UQ provides an associated measure of confidence. LANL has a rich history of applying UQ to diverse areas such as weapons physics, cosmology, and hydrology, and we are currently extending this set to include energy, environmental management, climate, and nuclear forensics. In the coming years we will extend the applicability of this framework to problems beyond the physical systems to which UQ was originally applied. New application domains include, among others, cybersecurity, remote sensing, materials science, and education. UQ is an integral part of modeling and simulation work at LANL.

Milestones FY14

- Identify existing UQ capability across LANL programs and organizations
- Conduct a Deep Dive to identify strategic directions for further development of LANL UQ capability
- Develop a plan for LANL UQ coordination
- Initiate steps to communicate UQ concepts, methodologies, and results to scientific staff and decision makers

Milestones FY15

- Carry out in-depth technical inventories of existing capabilities and technical challenges
- Identify gaps in knowledge and opportunities for cross-fertilization
- Develop an institutional strategy for ongoing development and utilization of UQ capability to support institutional missions
- Determine appropriate and useful methodology for obtaining uncertainty estimates associated with physical system calculations as well as functionals and other mathematical constructs
- Determine how uncertainties evolve from a calibration point of view to assess uncertainty present in predictions from calibrated models
- Develop a methodology for defensible extrapolation of calibrated models to regions beyond existing data that addresses the associated uncertainty
- Develop approaches for assessing models of nonphysical phenomena, addressing associated data types and structures

Milestones FY16 and Beyond

- Develop a scientific approach for modeling that leverages expertise in uncertainty quantification and computation to efficiently generate and analyze data, thus accelerating scientific discovery.
    - Understand the relative contributions from different types of uncertainties
    - Demonstrate the use of experimental planning and resource allocation to optimize predictive capacity
    - Incorporate a multi-scale framework into the UQ paradigm
- Initiate strategic partnerships with academic collaborators and other laboratories to strengthen UQ capabilities
- Identify and pursue high-quality funding opportunities
- Provide a coherent, Laboratory-wide effort in UQ applied to a broad set of problems

Stretch Goal

- Develop a nationally recognized UQ program and assume a leadership role in UQ within the DOE complex

Required Resources

- Support for exploratory activities, plan development, and program development
- Management support to champion this effort and facilitate removal of organizational barriers
- Access to existing technical staff and opportunities for staff recruitment

External Partners

- Academic collaborators
- Multilab programmatic partners
- PSAAP centers

Existing Efforts

- Programmatic efforts in nuclear weapons, energy, environment, and climate
- Collaborative research efforts with academic and government laboratory partners
- Participation in national committees, conferences, scientific publications, and planning activities

Theme: Visualization and Data Analytics

Scientific Visualization is an essential tool for understanding the vast quantities of large-scale, time-dependent data produced from high performance computer simulations. While interaction is recognized as a key feature for useful exploration of scientific data, sufficient speed for interaction is impossible on these large data sets using commercially available visualization software and algorithms. Therefore, an extensive research program is required to meet future requirements. The nature of the required research spans the areas of traditional computer graphics, scientific visualization and computer systems software.

LANL's current deliverables in this area focus on delivering scalable in-situ analysis R&D through several complementary technologies, for a specific set of partners. Building on the open source Catalyst and VTK-M capabilities, we will deliver advanced data reduction, sampling and visualization tools to help scientists analyze specific real world problems.

Milestones FY14

- Deliver in-situ analysis capabilities to ASC customers for ensemble analysis. This is a partnership between LANL analysts and data experts intended to impact the quality and efficiency of complex ensemble analysis. This work is a follow-on to the group's FY13ASC Level II Milestone.
- Deliver prototype R&D capabilities to ASCR for image database analytics, statistical sampling, and interactive interfaces for large data. Included in this are published papers, open source code, and working prototypes of the technologies.
- Deliver portable accelerated halo finding capability that works in-situ with HACC cosmology code for SciDAC SDAV project.

Milestones FY15

- Deliver robust in-situ analysis capabilities to ASC customers that enhance the collaborative analysis workflows developed in FY14
- Advance the capabilities of image based analysis, statistical sampling for extreme datasets, and web-based inline UIs for ASCR research customers

Capability: Computational Co-design

The primary challenge to computational physics and applied mathematics over the next decade will be a revolution in computing technology. The first instantiation of this was the Roadrunner supercomputer, the first system in the world to achieve a petaFLOP/s of sustained performance. The Roadrunner system had a variety of architectural features that underscore where computing is heading, including accelerators of a different technology design than the base system (heterogeneity), different and complicated memory hierarchies, multiple networks, multiple compilers, and other complexities that required a complete re-design of application codes and algorithms for the architecture.

As the nation drives toward exascale, computing resiliency and fault tolerance will also become issues that applications must face, demanding new strategies that go beyond the current checkpoint and restart methods employed today. In future Exascale systems, applications will have to tolerate node-level failures that occur at a frequency measured in minutes, and will require scalable recovery solutions that do not rely on globally accessible checkpoint data. This will drive significant changes in the programming models that will run on these systems, and require hardware and software capabilities that do not currently exist.

It is obvious that in order to exploit emerging architectures, computer scientists and computational physicists will be required to work together even more closely and develop long-standing collaborations in order to effectively map new computational physics algorithms onto new computer architectures, and to generate new computing architecture design ideas based on computational algorithm requirements. We have embarked on a serious and well-funded co-design effort in which methods developers, computer scientists, and computational physicists work together with industrial partners designing and creating new computing technologies to *co-design* the technologies and applications. This is not a new approach, as space-based computational assets have been designed in this manner for some time. However, it is the scale of effort, technical scope, and broad application approach that are different. The goals for computational co-design are:

- Establish LANL's leadership role in computational co-design
- Ensure that co-design is an integral part of the MaRIE signature facility
- Develop a critical mass of technical staff with co-design expertise
- Secure a broad portfolio of externally funded programs in computational co-design
- Develop a generalized process of co-design to bind together experiments, applications, and technology and create a truly integrated science and technology enterprise at LANL

Theme: Applied Computer Science

One event that upsets the longevity model of scientific software is a disruptive change in computing technology. These changes are generally driven by hardware cost or manufacturing constraints. In scientific computing, there have been two previous technology upsets: vector computing and massively parallel computing. Both of these events required that scientific applications be overhauled to obtain the performance gains promised by the new technology. We are currently in the midst of a third technology upset, namely the advent of heterogeneous multicore computing. "Multicore" indicates a move to having many processors physically packaged together and "heterogeneous" implies that those processors have differing performance and functional characteristics.

We believe this technology upset will also cause applications to be significantly re-designed. Scientific computing is in the unfortunate state of knowing that the hardware is already changing, but not knowing how to extract benefit from those changes. The only way to take advantage of the current upheaval in hardware is the collaborative development or co-design of hardware, system software, algorithms, and applications with each informing the other so we achieve the best performance for a given set of constraints in price, performance, power, and resilience.

This rapidly changing technology, already underway, forces a pragmatic approach to co-design through the formation, organization, and management of multi-disciplinary teams working on a single problem of common interest. Our ultimate goal is to make computational scientists more productive, regardless of system size. Along the way to this goal, there are many methods and technologies that can be transferred from forward-looking research to a productive application development. This is achieved through a balance between R&D, knowledge transfer, and community building to achieve its long-term goals while being relevant and productive in the near-term.

Current Co-Design Projects

LANL's strategic investment in co-design supports multiple projects. Below we describe some of these current projects to give some flavor of the breadth of co-design technical areas covered.

ExMatEx (Exascale Co-Design Center for Material in Extreme Conditions), is an ASCR-funded project (one of three co-design centers) that focuses on scale-bridging materials modeling. This is primarily a computer science project driven by a specific application focus. The four main components of the project are: proxy applications, computer science, multi-scale algorithms, and validation and verification. Multi-disciplinary teams focus on the iterative design and constructions of scalable proxy applications that explore the problem domain. Innovative computer science practices in tools, simulators, programming models, and run-time systems are used to enable efficient, dynamic, concurrent, and asynchronous material modeling applications.

CoCoMANS (Computational Co-Design for Materials and the Natural Sciences), is an institutionally funded LDRD project that focuses on algorithmic method for multi-scale simulations of plasmas and ocean circulation. In addition to algorithmic innovations, co-design processes are used to optimize these algorithms with today's (and future) heterogeneous computing systems.

Co-design projects within ASC are working to evolve our weapons codes to effectively execute on these new heterogeneous architectures. As in other co-design projects (here and elsewhere) the co-design efforts revolve around the development of representative proxy applications. We are working closely with vendors engaged in the DoE FastForward project to optimize these proxy applications to their future hardware. Additionally, we are working to bring modern computer science principles into the process through active co-design engagement between computer and domain scientists,each group alone is insufficient to develop a sophisticated code that is both portable and performs well. Our goal is to develop co-design techniques and principles that can be applied to new code efforts within the weapons program.

Milestones FY14

- Affect the design of microprocessors or other technologies that end up in pre-exascale and exascale systems, through the process of co-design
- This work will continue with the DoE DesignForward project (follow-on to the FastForward project)
- Integrate Exa14 class activities directly into weapons code development projects by direct inclusion of computer science participants into code projects
- Develop domain specific language compiler that maps data translation codes directly to FPGA so developers do not need to program that device at a low level (currently customary)
- Fully merge mini-app effort into XCP code project by direct inclusion of computer science participants into code projects

Milestones FY15 and Beyond

- Demonstrate scale-bridging applications at 10+ petaflops
- Integrate Exa14 class activities directly into weapons code development projects by direct inclusion of computer science participants into code projects
- Develop a domain specific language compiler that maps data translation codes directly to FPGA so developers do not need to program that device at a low level (currently customary)
- Fully merge mini-app effort into XCP code project by direct inclusion of computer science participants into code projects

External Partners

- LLNL
- SNL
- ORNL
- Stanford
- Caltech

Vendor Interactions

- nVidia
- Intel
- AMD
- ARM
- IBM

Theme: Computational Mathematics

Computational mathematics at LANL includes the development and application of numerical and symbolic algorithms to the computational solution of scientific problems. The quality of numerical methods and algorithms as well as critical mass of researchers working in computational mathematics is extremely important for success of multiple capabilities of fundamental significance at LANL. Research in computational mathematics must be driven by the requirements and needs of other capabilities described in the implementation plan for the IS&T pillar strategy, and in particular, other elements of the co-design capability.

FY14 Milestones

- Identify gaps in capabilities for current important applications including those in XCP, XTD, CCS, T, and EES divisions
- Identify key integration points for existing capabilities
- Assess workforce capabilities for future code development and identify critical gaps
- Hold internal (LANL) workshop on Computational Mathematics (focusing on determination of capabilities, needs and customers)
- Work with XCP Leadership to include elements of Computational Mathematics in PSP
- Work with LANL Fellows to promote Computational Mathematics at LANL
- Establish communication with ASCR (DOE Office of Science) Exascale Mathematics Working Group – LANL Contact is Luis Chacon

FY15 and Beyond Milestones

- Plan strategic institutional investments (LDRD) to start filling identified gaps
- Publish capability gaps to specific line organizations for strategic hiring
- Create a suite of institutionally and programmatically funded projects to target identified areas of opportunity
- Create a funding envelope for innovation to constantly refresh the numerical methods in preparation for exascale systems
- Work with ASCR DOE/SC to define niche in Computational Mathematics for LANL
- Create a strategic plan for development of Computational Mathematics Capability at LANL

Existing Efforts

- ASC projects: Code Projects, Hydro, Transport
- Office of Science climate projects
- Advanced Scientific Computing Research (ASCR) DOE Office of Science Project on Mimetic Methods for PDEs
- LANL LDRD DR Project on "New Generation Hydro Methods"
- Advanced Simulation Capability for Environmental Management (ASCEM)
- Consortium for Advanced Simulation of Light Water Reactors (CASL)*External Partners*
- UCSD, FSU, UCD, UOH TAMU, and other universities
- US National Labs – LLNL, SNL, ORNL
- CEA – France
- AWE – UK
- Russian Labs

Required Resources

- Support from XCP, CCS, and T management and ASC management
- Consistent institutional investment strategy that leads to a stable funding portfolio
- A workforce ready for the computational and technological challenges

Theme: Computational Physics

Computational physics at LANL covers a broad range of disciplines and scales from specialized tools and methods focusing on single physics at the atomistic or even quantum scale to large multi-physics integrated capabilities at the continuum scale and even engineering codes built upon a high degree of empiricism. It is desired and necessary that the efforts spanning this spectrum are coupled in such a way that an overall capability at LANL is available to address problems in the national interest. A dominant driver is the pragmatic need for predictive large-scale multi-physics capabilities for climate and weather simulation, energy systems simulation (nuclear fission reactors, combustion, etc.), and nuclear weapons performance simulation. These capabilities rely or attempt to rely upon the lower scale efforts to inform reduced order models or even provide a basis for subscale resolution within the larger-scale codes. They also require coupling between the various physics components. Most LANL simulation tools employ traditional low-order operator splitting to couple multiple physical phenomena within an integrated simulation. While these methods are straightforward to implement, there is limited understanding of their stability and accuracy constraints. Frequently, such simple methods lead to significant time step restrictions for stability.

In addition, the new multi-core architectures we face require computational co-design wherein the development of methods and algorithms and hardware design, to the extent possible, are tightly coupled, with neither considered independent of the other. This new paradigm creates an opportunity to truly inline the ability to bridge scales of interest and to utilize this additional compute power to greatly improve multi-physics coupling. LANL is actively addressing this triad through a number of co-design activities. One is the LDRD DR project *Computational Co-design for Multi-scale Applications in the Natural Sciences (CoCoMANS)*. This project will develop new "applications-based," self-consistent, two-way, scale-bridging methods that have broad applicability to the targeted science In addition this project will map well to emerging heterogeneous computing models (while concurrently guiding hardware and software to maturity), and provide the algorithmic acceleration necessary to probe new scientific challenges at unprecedented scales. Another effort is the *Eulerian Applications Project – Future Architectures (EPA-FA)* that is developing a light weight set of computational tools and resources to abstract physics from data access. The intent is to provide a transition plan for many of our large-scale multiphysics codes for climate, energy, and nuclear weapons to multi-core platforms without loss of the institutional knowledge built into these resources.

This theme area is meant to collect such activities to the extent they are not represented elsewhere in this Plan so that important integrated multi-physics or unit-physics applications are appropriately represented. There will be overlaps in these activities with other themes in the implementation strategy as computational physics codes (even the unit-physics variety) are themselves a fundamental integrating element. Topics include:

- Computational fluid dynamics methods (mesh and particle approaches) and integrated codes
- Neutral and charged particle transport methods (Monte Carlo, deterministic, hybrid) and integrated codes
- Macro- and meso-scale computational solid dynamics and thermal hydraulics
- Particle-in-cell methods for kinetic plasma simulation
- Particle-in-cell and other methods for kinetic plasma and turbulence simulations
- Molecular dynamics for plasma simulation especially in the dense non-ideal regime; advances in methods and computing hardware make this a remarkable tool to address questions not previously accessible without a high degree of approximation
- Reduced order models from microphysics simulations and sub-grid models built upon microphysics capabilities
- Computational magnetohydrodynamics
- Phase field methods
- Computational manufacturing applications
- Integrated nuclear weapons applications
- Integrated climate and ocean modeling applications

Milestones FY14

- Identify current gaps in sub-scale models and ability of modern architectures to allow direct integration of sub-scale physics
- Determine appropriate integration pathway for subscale models and methods into multi-physics capabilities
- Create a light-weight suite of computational tools for a computational backplane resilient to technology change and broadly accessible to the scientific and industrial communities
- Point these efforts at emerging needs and growth areas

Milestones FY15

- Identify weak leveraging between efforts at different scales across the laboratory
- Identify key integration points for existing capabilities that are currently missing (e.g., the integration of UQ in emerging or existing application areas)
- Plan strategic institutional investments (LDRD) to start filling identified gaps
- Reach out to external partnerships in specific areas
- Strategy for retention and recruiting

Milestones FY16 and Beyond

- Create a suite of flexible applications on a variety of computing hardware at multiple scales
- Create a funding envelope for innovation to constantly refresh the technology, algorithms, methods, and efficacy of a core set of integrated codes in preparation for exascale systems

Stretch Goal

- Acceptance of flexible computational backplane to by Weapons, Global Security, and Office of Science Programs
- Creation of a Computational Physics Center focused on problems of interest to the core mission of the laboratory to attract and retain the best and the brightest in support of our core missions in Weapons, Climate, and Energy

Required Resources

- Consistent institutional investment strategy that leads to a stable funding portfolio
- A workforce ready for the computational and technological challenges
- A suite of talented software developers

Existing Efforts

- ASC weapons code projects
- Office of Science climate projects
- Carbon Capture Simulation Initiative (CCSI)
- ASCEM
- CASL
- Other related work for others projects requiring these technical capabilities

Theme: Novel Computing

This theme is centered on groundbreaking new computing paradigms, devices, and disruptive innovations in all areas of information technology for the next 5-20 years. The overarching goal is threefold: 1) ensure the progress of silicon-based computers, 2) ensure continuing progress beyond Moore's Law and beyond silicon-based computers, and 3) enter into unimagined or presently unreachable applications (e.g., in terms of scale) and application domains.

Silicon-based architectures continue to be developed and extended. As these systems become more capable and more powerful we need to see if they can be adapted to the unique computational needs of the Laboratory and beyond. Microprocessors are becoming ubiquitous in electronics, cell phones, tablets, and music players. These systems are designed for lower power and no cooling, yet they scale to the billions of interoperable nodes. New silicon devices are being developed that may be able to support Lab missions.

With the ever-decreasing node dimensions of today's computer chips (based on von Neumann architecture) reaching physical and practical limits, questions arise about what will come next and how we, as a laboratory, can push the state of-the-art computing architectures and remain on the cutting edge. The Microelectronics industry that has been driving computing has reached two limits: 1). device scaling (smaller dimensions will transition into the quantum regime) and 2). power issues (exascale computers will require at least 20-30MW of power). Additionally, the von Neumann model is one of load-compute-store and gives rise to moving the data to the process, and we can't afford that abstract model due to data movement costs (power/latency). Moreover, conventional computational systems are not well suited for certain applications: von Neumann systems perform exact (deterministic) computations extremely well while humans excel at abstraction and prediction based on prior experiences (an aspect of a probabilistic approach). While we do not quite understand the neural wiring of the human brain, there must be a synergy (computational co-design) to enable a level of computing that combines the advantages of a quantum, neuromimetic, traditional and other approaches.

In answering the challenge of roadmapping the future of computing, both hardware and software must be taken into consideration: developing new code without new hardware architecture is unlikely to produce breakthroughs, as is producing new hardware without developing "operating systems" that can exploit its capabilities.

Milestones FY14

- Enumerate possibilities for novel computing
- Determine what areas of novel computing are ripe at LANL to be further developed
- Identify staff interested in developing novel software and hardware

Stretch Goals

- Attract external funding for exploring possibilities in novel computing

Required Resources

- Program Development to support staff in shaping LANL's novel computing portfolio
- Funding for travel to customer sites
- External Partners
- Sandia National Laboratories
- Portland State University
- Commercial entities as needed

Theme: Simulation Science

The simulation science theme is centered on evolving simulation techniques, where we define simulation as a model of a real system and its processes that attempts to capture the temporal features of such a system. As opposed to other modeling paradigms, simulation always advances time. While this definition would encompass traditional physics-based simulations (such as MD), which are treated in other themes, we restrict the simulation science theme to applications that have an engineering or social behavior component. Distributed discrete event simulation (DDES) is a simulation technique that uses discrete events as a main driving force to simulate the evolution of a complex system. The notion of discrete event simulation and the theoretical underpinnings were established in the 1970s and 80s. Discrete event simulation is used predominantly for engineered systems that follow a set of prescribed rules, such as the Internet and communication networks in general, supercomputers (relevant for the optimization principles in the co-design subtheme), vehicular transportation networks with discrete traffic-light logic, epidemic modeling, or war games simulation. Any other process with discrete events can be modeled using discrete event simulation, most notably biological and chemical processes. Discrete event simulations also play a role in composite systems where physical systems interact with social systems, such as in modeling global warming. The physical processes of greenhouse gas generation, atmospheric mixing, oceanic interactions, and terrestrial uptake are relatively well understood. However, the social processes such as agriculture, land use, electric power generation, transportation systems, and economic choices at multiple social levels are much less well understood. Agent's attempt to approximate the perceptions and decisions of humans despite acknowledged difficulty in predicting human behavior.

We distinguish four subthemes: 1) Cyber Systems and Defense, 2) Socio-technical systems, 3) HPC Performance Prediction, and 4) Foundations of Simulation Science.

Cyber Systems and Defense

Modeling and Simulation (M&S) is regarded by many sponsors (including DHS and DoD) to be a main pillar of any cyber security strategy. Cyber models predict behavior of cyber components for novel offense and defense strategies without having to start an actual cyber war. In addition cyber models can be used for training purposes. The current LANL portfolio of cyber models include:

- MIITS (Multi-scale Integrated Information and Telecommunications Systems), a scalable packet-level communication network simulator that models heterogeneous networks;
- CyberSim, a tool with graphical interface that simulates the spread of a cyber exploit through the computational fabric based on software stacks and user behavior; it also includes social network behavior;
- BotSim is a scalable tool for testing advanced botnet schemes in an offense and defense setting, where a botnet is a collection of computers that are controlled by a single entity unknown to the legitimate owners of these computers; and
- GIAC, a long-running mature battlefield simulation project. The main technical challenge in this area is in developing more accurate models of user behavior, where we propose to distinguish attackers, defenders, and bystanders, which are regular computer users. The main challenge from a funding perspective is to find a committed sponsor for a signature project.

Milestones FY14

- Develop MIITS-like capabilities in new Python-based SimX package
- Develop more accurate defender and attacker models potentially integrated into a game-theoretic framework

Milestones FY15 and beyond

- Update behavior models and grow catalog of modeled technologies; the communication area changes quickly (see for example the move from Facebook to Twitter); we need a committed sponsor to stay up-do-date constantly

Socio-technical Systems

LANL has a long history of developing simulation models of socio-technical systems. Socio-technical system models include the NISAC-developed and now NIH-sponsored EpiSims model for disease spread; EpiSims relies on the Dynamic Activities model that simulates agents who plan and optimize their daily schedules. FastTrans is a large-scale vehicular traffic simulator that grew out of the TRANSIMS project. EduSim is an agent-based model of students that make their way through the education system. Themis-1 is an economic behavior model that currently focuses on monetary policy. We have purposely let socio-technical systems grow in many directions. Only the EpiSims/Dynamic Activities project has had a reliable funding stream. Given the ever-increasing importance of social phenomena, the goal is to find an additional sponsor to turn one of these tools into a long-term viable development. While all application domains are important, the economic modeling of Themis-1 has attracted the most interest from internal sources.

Milestones FY14

- Develop Themis-1 as a scalable Python-based SimX application
- Identify and develop sponsor for extensions of Themis
- Find follow-on sponsor for EpiSims and Dynamic activities work

Milestones FY15 and beyond

- Develop additional models on an as needed basis

HPC Performance Prediction

Computational Codesign projects, such as the LDRD DR project titled “Optimization Principles for Co-design applied to Molecular Dynamics”, have performance prediction of abstracted software on abstracted hardware architectures at their core. Scalable and accurate performance models for supercomputers and novel architectures are a stated need (by ASC and ASCR) in any progress towards next-generation systems. Leveraging LDRD and ASC funding, the LANL portfolio in HPC performance prediction tools include 1) FileSim, a simulator of distributed file systems, 2) AMDSim, the core simulator of the LDRD DR project of simulating an application code (AMD) on novel hardware, and 3) ASET (Architecture Simulation and Exploration Tool), a near-cycle accurate hardware simulator.

Milestones FY14

- Demonstration of the closure of an optimization loop over an algorithmic space representing the MD problem using AMDSim and possibly ASET
- Burst-buffers modeled in FileSim to help guide trinity decisions
- AMDSim/ASET: Demonstration of the optimization loop–the iteration of HW/SW pair selection and performance prediction–over non-trivial HW and SW spaces for a scientific computing application area of interest, in our case a particular class of MD calculations
- Progressive development and refinement of above and their incorporation into the optimization loop; by 2014, Q3, the end of the project, we aim to deliver an optimization loop methodology, with particular emphasis on "non-trivial" that is, exploring HW and SW spaces where local and global maxima of given objective functions are not obvious, even to human experts
- Merge the FileSim code into the AMDSim/ASET code base to become the basis of a comprehensive performance prediction suite

Milestones FY15 and beyond

- Add more applications to application simulator concept of AMDSim

Foundations of Simulation Science

As pointed out by the CPAM review committee, LANL needs to invest in the foundations of simulation science. Scalability, the traditional strength of LANL simulation work, should be reinvigorated coupled with fast prototyping and easier use. The original SimCore framework has been transformed into an easier to use scalable simulation engine now written in Python to allow for a larger user base. Features such as process-orientation have been added.

Milestones FY14

- Conduct scaling tests with SimX on benchmark runs, in the tradition of phold
- Establish formal partnerships with leading academics in simulation, including David Nicol and Jason Liu

Milestones FY15 and beyond

- Grow user-base of SimX within government institutions

Capability: Complex Networks

Networks provide a natural framework for describing complex systems in terms of interdependent subsystems and components. Complex systems and their network representations arise in many scientific and national security mission areas at LANL: cyber systems, national infrastructure, biological systems, real and artificial neural systems, classical and quantum communication systems, systems of networked sensors, social networks, terrorist networks, and energy systems such as the smart grid. Across this spectrum of complex networks are common R&D frontiers that include discrete mathematics and graph-based methods, as well as specific algorithmic and computing architectures.

The explosion of research in network science is providing many new approaches and techniques for analyzing and predicting the structure and dynamics of socio-technological systems such as the spreading of epidemics and communication patterns in social networks. The ability to easily collect and store large amounts of data has spurred the progress of research even more as the rush to find hidden patterns and relationships has hit everywhere from e-commerce businesses to the intelligence community.

At the heart of these problems are mathematical models and algorithms. Our set of projects aims at constructing network models that are relevant and mathematically analyzable, designing network metrics and scalable algorithms, and optimizing network topologies and performance metrics. A long-term goal is to find relationships between dynamic network processes and network structure. We apply these models and metrics to sociological and technological applications such as securing cybersecurity networks, measuring the impact of scholarly publications, defending against nuclear smuggling, and recovery of systems after weapons of mass destruction (WMD) attacks. The goals for complex network science are as follows:

- Establish LANL's leadership role across a spectrum of application areas in complex networks science including cybersecurity, smart-grid systems, and metagenomic systems
- Establish LANL's leadership role in the DOE and government for cognitive neuroscience for national security, and position LANL to compete for external program opportunities requiring an advanced synthetic cognition capability
- Achieve visual synthetic cognition with an artificial neural network, as measured by the ability to outperform a human in certain recognition tasks
- Demonstrate quantum-enabled secure communication on a networked system of users

Theme: Complex Process Networks

Networks of Complex Processes

In many situations, understanding the topology of a complex network is not enough to reveal what it does, or why it functions in a particular way. It is necessary to look more deeply at what processes are embedded in the network to see how the functioning of one process enables or disables the functioning of another, what one process produces that another consumes, and how the strength and timing of their interactions influence the behavior of the system overall.

Complex process networks are ubiquitous in science and technology, from the signals that propagate information from the receptors of a cell's membrane to the critical, interwoven infrastructures that make modern civilization possible. The overarching goal of this theme is to help scientists and engineers characterize and analyze complex process networks based on ideas and principles that cut across a wide range of knowledge domains.

Because our interest in analyzing networked processes is very general compared to those who currently analyze large, complex process networks within specialty fields, we believe it is important to foster a community of interest. We believe this involves two strategies:

- Suggest the value of pursuing a general approach to analyzing complex process networks by encouraging the application of general principles to other capability and application areas already identified within the IS&T pillar
- Promote the development of general tools for characterizing and analyzing complex process networks

Connection to IS&T Capability Areas

We suggest several capability areas already identified within the IS&T pillar that we believe would benefit from enhanced capabilities in complex process network analysis.

Data Science @ Scale

Here the basic theme would be *characterizing and analyzing computational workflows*. Whether the goal is to use big data to develop statistical models, as it is in *Machine Learning*, or the aim is exploiting signatures in large datasets as it is in *D2D*, a general approach to dealing with large computational workflows would constitute a powerful conceptual framework.

Co-Design

Computational Physics and *Simulation Science* are two capability areas that in the large arena share the same challenges of coordinating computational processes into a coherent workflow. As in the case of Data Science @ Scale, the benefit of a larger perspective would serve here as well.

Complex Networks

Other sub-themes within the complex network capability area would also be served by this theme. Specifically, *Cyber Security* depends on analyzing large networks of computer assets, not only with regard to connectivity, but in terms of functional interdependence, system capacities, and timing. These are core aspects of process networks that a general approach to analysis would bring into play. Developing capabilities in *Synthetic Cognition* also requires the application of such principles.

Connection to IS&T Application Areas

We believe that several application areas identified within the IS&T pillar are critical to the laboratory's mission. These areas would benefit from the emphasis on a general approach, to process network analysis, that also promotes the development of scalable sharable analysis tools:

- Bioinformatics and systems biology
- Nuclear informatics and situational awareness
- Climate informatics and climate systems science
- Smart Grid and other critical infrastructure

Tools for Complex Process Networks

A modest effort, using institutional funds in the form of technology maturation grants, has been underway since 2012 to produce a database engine and analysis tool based on the general ideas described above. The project's working title is the *Functional Reasoning Engine* or *FRE*. Its intent is similar to that of graph databases used to characterize and analyze large network graphs, but adds an additional layer of semantics useful for representing embedded processes. A prototype implementation exists, and a more robust version is in progress.

Theme: Cybersecurity

Cybersecurity is the still-emerging concern of the security of complex systems and networks of distributed computers, as well as the information and infrastructure controlled by them. Cybersecurity is widely regarded as one of the newest and largest security concerns for government, industry, and society. Like other types of security, the field inherently includes both offensive and defensive issues. Cybersecurity includes academically-recognized specialties such as cryptography, computer security, network security, information security, computer and network forensics, and intrusion detection, as well as government specialties such as computer operations, exploitation, information operations, and reverse engineering. The importance of cyber security increases year-by-year in the eyes of the public and government. The looming threat of cyber warfare is increasingly acknowledged by the President and the Secretary of Defense. Cyber infrastructure has proven to be widely vulnerable to exploitation, and exploitation is easily leveraged into disruptive attacks. While operational cybersecurity funding is increasing in government, long-term R&D is limited largely to a $600M portfolio overseen by OSTP and executed by (in decreasing order of spending), DARPA, DOD, NSF, DHS, NIST, and DOE OE. Efforts such as the Comprehensive National Cyber Initiative and the Cyber Leap Year have thus far failed to drastically reduce national security risks. The country needs game-changing approaches to cybersecurity that fundamentally shift the balance between attacker and defender.

LANL's strategy for developing cybersecurity science & technology is stewarded by the Cyber S&T Program Council led by the Advanced Computing Solutions Office in ADTSC. Our strategy is to leverage and sustain LANL's broader capabilities—many of them IS&T—to make impactful contributions to cybersecurity. Our goal is to achieve science-based, game-changing impacts. We leverage capabilities such as statistics, machine learning, complex networks and dynamic graphs, quantum, graph interdiction, computer science, discrete-event simulation, and big data. Investments seeded by LDRD have put LANL in a leadership position with respect to detection in dynamic graphs and quantum-based security. In the near-term, we are seeking to leverage the more recent dynamic graph work into sustained R&D funding by other agencies. Meanwhile, we are seeking to build new leadership areas in predictive vulnerability analysis, quantified resilience, and short-timescale cyber engagements. Due primarily to the lack of an NNSA cybersecurity R&D program, we are heavily focused on developing sustainable funding relationships with other federal agencies, DARPA, DOD, and DHS S&T in particular. However, the tight time-scales in cybersecurity drive potential sponsors to push for increasingly mature, deployable technologies and we are actively pursuing commercial partnerships to accelerate that technology impact.

Milestones FY14

- Capture of FY15 LDRD investments for cyber security projects
- Integrate cybersecurity S&T strategy with Strategic Outcomes office planning & initiatives
- Support the establishment of a base-funded DOE/NNSA Cyber Sciences program
- Capture DARPA seedling in cybersecurity
- Obtain and enlarge a second-year of funding from a strategic WFO partner
- Highlight cybersecurity S&T web site and accomplishments in LANL's main web presence
- Co-host 2nd International Workshop on Interdisciplinary Cybersecurity

Milestones FY15 & Beyond

- Capture DHS S&T BAA funding as the prime contractor
- Sustain cybersecurity as a strategic application area for LDRD research funding
- Re-start the Cyber Defenders student program

Required Resources

- Operational network environment for testing new sensing and response technologies
- Sustained real-time data access & curation of datasets by OCIO programs and ADIT
- LDRD strategic investment support
- Cyber defenders program for students
- University institutes to maintain a staffing pipeline
- TT support for the Cyber Resilience Center (name TBD) engagement with industry & academia and other partnerships
- Institutional computing support for time traditional HPC resources and/or HPC-managed data-intensive platforms (with sufficient data security for institutional data)
- Sustained think-ahead funding to anticipate threats and seed reactions to them
- Participation by program offices and inclusion in high-level LANL overviews as a strategic mission area
- Sustained Program Development funding to enhance and maintain LANL's visibility in cybersecurity and to build new relationships

External Partners

- Tri-lab cybersecurity leadership
- Imperial College London
- Strategic federal agency partners (DARPA, DHS, DOD)
- Commercialization partners (in negotiation)
- Missouri S&T University
- New Mexico Tech

Theme: Quantum Information Science

This theme is centered on research that exploits the power of quantum mechanics to enable new techniques to acquire, process, and protect information. A strong capability in Quantum Information Science (QIS) has been one of the goals of the IS&T pillar over the past few years. QIS is widely recognized as a disruptive technology with the possibility of rapidly changing the technology landscape in areas that will continue to be important to the Laboratory such as information security, simulation and design of materials, and state-of-the-art sensing. LANL research in QIS has three sub-areas: 1) quantum metrology and other quantum-enabled technologies, 2) quantum algorithms and architectures (quantum information theory, superconducting, quantum-dot and semiconductor qubits), and 3) quantum many-body physics and cold-atom physics (decoherence, foundations, quantum phase transitions, cold atoms, casimir physics). Two major goals dominate our strategy for the next few years: (1) develop the new QIS capability in the area of quantum metrology and sensing, exploiting both the enhanced sensitivity enabled by quantum correlations and the potential advantages in size-weight-power-cost of quantum systems over conventional sensing technologies, (2) renew our engagement with the national quantum computing program.

Milestones and partnerships that apply to all sub-areas of the QIS theme include:

Milestones FY14

- Successful coordination of quantum research involving postdocs, and operation of the Quantum Lunch, Distinguished Quantum Lecturer, and visitor programs by the Center for Nonlinear Studies (CNLS)
- Establish the QIS Theme of the Information Science and Technology Institute as a focal point for the LANL QIS community, and initiate monthly meetings of technical staff and management interested in QIS
- Improve coupling of QIS basic research to program offices by (a) identifying customers for LANL QIS and, (b) by arranging one-on-one meetings with relevant LANL program managers
- Cultivate connections with engineering capabilities in PADGS and ADE, allowing LANL to deliver complete QIS systems
- Develop connections between QIS and the "Science of Signatures" pillar
- Nurture QIS connections to UNM Center for Quantum Information and Control and SNL, including institutional participation in the Southwest Quantum Information and Technology (SQuInT) network

Milestones FY14 and beyond

- Conduct workshops for external sponsors, such as the November 2010 "Quantum Technology Applications Workshop"
- Continue execution of the quantum strategy to build a sustainable QIS portfolio
- Develop a sustainable QIS portfolio with a mix of basic and applied research supported by both internal, Weapons Facility Operations, and programmatic funds

Required Resources

- Support for basic science research in QIS sub-areas, providing base funding while efforts ramp up to the point where external sponsors are willing to pick them up

External Partners

- UNM's Center for Quantum Information and Control (LANL technical staff serve as adjunct faculty, UNM graduate students perform PhD project research at LANL, and LANL is an institutional member of SQuInT)

Quantum Metrology and Quantum-Enabled Technologies

Milestones FY14

- Collaboration between physicists and engineers delivers complete atomic magnetometer system
- Proof-of-principle rotation sensing with Atom-SQUID
- Develop techniques and technologies to harvest entropy from quantum states and systems; extract and exploit quantum randomness for applications ranging from fundamental physics to information security

Milestones FY15 and Beyond

- Advance state-of-the-art in quantum-enhanced communication science and technology; develop strategies and techniques to overcome longstanding limitations in communication rate, range, and scalability
- Demonstrate matter wave circuits for sensing and other applications
- Obtain substantial external funding for a program in quantum metrology

Required Resources

- Assistance in making connections between relevant technical areas at LANL and in assisting with program development

Existing Efforts

- Projects funded by LDRD, DARPA, and IC

Quantum Algorithms and Architectures

Milestones FY14

- Further develop LANL quantum algorithm theory
- Engage Materials Theorists to improve the state of the art in semiconductor qubits
- Build experimental capabilities in topological quantum computing, semiconductor qubits, and superconducting qubits
- Develop new materials for quantum architectures (using topological protection)
- Develop methods for detecting and avoiding photon number splitting attacks in quantum cryptography, under the bounded storage model

- Develop fast streaming algorithms for independent quantum random bit generation
- Construct novel quantum algorithmic techniques for fast signal processing using wavelets
- Develop and apply quantum methods for computing ground state properties that are also suitable for the simulation of quantum circuits
- Develop and apply quantum methods to simulate stochastic differential equations
- Develop fast quantum algorithms for linear algebra problems
- Develop quantum methods for adiabatic state transformations that are faster and more robust than adiabatic quantum computing

External Partners

- SNL, UNM, University of Pittsburgh, UC Irvine

Existing Efforts

- Research is funded by LANL LDRD, SNL-LDRD, AFOSR, and NSF, with potential future funding from NSA, IARPA, and DARPA.

Quantum Many-Body Physics and Cold-Atom Physics

Milestones FY14 and beyond

- Investigate superpositions of collective degrees of freedom in interacting many-body quantum systems
- Study how many-body systems act as sources of decoherence that eradicate non-local superpositions
- Develop the connections between quantum computing and quantum phase transition physics in condensed matter systems

Existing Efforts

- Research is funded by LDRD

Theme: Synthetic Cognition

Understanding and emulating the function of the human brain is the overarching goal of the Synthetic Cognition initiative at LANL. This theme spans all of the current focus areas within the IS&T pillar: In neural systems the complex network topology not only mediates communication but also directly embodies the computation. Neural systems efficiently and effectively process and understand imagery, language, and a range of sensor data, giving neuromimetic systems important potential roles in big data analysis. Finally, neural systems employ unique data representations, computational primitives and algorithms that do not always map well onto conventional, general-purpose computing hardware. Thus computational co-design strategies are essential to realize the potential of neural–inspired computing to address difficult applications and to address emerging challenges of high performance computing.

The application of the cognitive and neural sciences to national security holds great promise for growth at LANL, as these areas address (and leverage) internal strategic goals and externally funded national security and defense needs. National security sponsors have been growing their portfolio of neuroscience projects, and this trend is continuing with the announcement of the national BRAIN initiative. Recent calls include novel biomedical imaging systems, bio-mimetic sensors and networks, advanced prosthetics, human-machine interfaces, and synthetic neuromorphic “brains” that approach the performance of biological system This is the beginning of a period of significant growth, driven by technical opportunity and a range of mission requirements.

Required Resources

- Institutional support for neural and cognitive sciences, computational neuroscience, and neuromimetic computing through the LDRD DR, ER, and Postdoctoral programs
- Continued support for the Synthetic Cognition theme to coordinate activities across organizations within LANL and to serve as a point of contact with the external community
- Program development efforts and support to respond to major national initiatives, including the BRAIN initiative recently announced by the Obama administration
- Strategic hires and support for postdoctoral fellows and early- to mid-career staff collaborations and alliances with academic and research centers and commercial entities

External Partners

- Revitalize historic neuroscience partnerships with external organizations, including the MIND Research Network and the Harvard/MIT/MGH Imaging Center
- Strengthen links to outstanding and relevant capabilities in experimental and computational neuroscience at UC campuses including Berkeley, LA, and San Diego
- Build partnerships with commercial and academic entities for neuroscience technology in national security applications, neural computation and neuromimetic system development

Existing Efforts

- Synthetic Cognition funded by LDRD and the DARPA Neovision2 project, both recently ended, produced a core of neural systems theory and cognitive experimental capability, establishing a foundation for LDRD projects in video processing and machine learning
- The Petavision project funded by NSF develops and applies HPC tools for neural simulation
- New projects from DARPA (UPSIDE) and NNSA address cognitive applications with neuromimetic algorithms and specialized computing architectures and hardware

Cognitive and Neural Science

The cognitive and neural sciences promise is to reshape the scientific and technological landscape of this century just as quantum physics, the human genome project, and digital computing redefined the last. Advances in experimental technologies and computational modeling capabilities set the stage for fundamental advances in our understanding of the architecture and dynamic function of the brain. Advances in materials, fabrication techniques, sensors and circuit design provide a path toward bio-cyber hybrid systems and neuromorphic systems. Potential applications range from biomedicine to national security and beyond.

LANL research strategies in cognitive and neural science have long focused on the development of analytical and measurement technologies to allow the study of large neural populations, recently adding capabilities for psychophysical study of perceptual and cognitive processes. This enterprise requires sophisticated, high-performance computational techniques to acquire, process, reconstruct, and analyze multiple, high-density data streams.

The role of the neurosciences to address national health issues will continue to increase in importance. The evolving understanding of brain function and continuing development of functional measurement technology will aid diagnosis, monitoring, and treatment of neurological and degenerative disease, mental illness, and traumatic brain injury. These efforts also provide the foundation of our efforts to identify functional principles of neural computation, and ultimately to mimic the function of neural systems. Such capabilities offer an important and unique role for LANL beyond the traditional sponsors and objectives of neuroscience research.

Milestones FY14

- Recruit new staff and revitalize programs focused on computational analysis, modeling, and integration of functional neuroimaging data
- Develop competitive proposals on basic neuroscience, brain imaging and neural interface technology, and strategic applications for NIH and Department of Defense (DoD) sponsors

Milestones FY15

- Develop instrumentation, data acquisition, processing, and analysis procedures to map neuroanatomy, connectivity, and brain activation dynamics in a single instrument
- Develop new capabilities for tissue impedance mapping and functional MRI at ultra-low field
- Implement advanced systems for brain-machine interface and sensory substitution

Milestones FY16 and Beyond

- Explore the proposed use of tissue impedance and/or direct detection of neural currents for dynamic measurement of neural activation in brain tissue
- Characterization of cortical source dynamics by MEG coupled with mapping of early visual areas via functional magnetic resonance imaging (fMRI), to characterize the perceptual role of synchronized oscillations within and between visual areas

Stretch Goals

- Develop and demonstrate fundamental new capabilities for human brain mapping, noninvasive and implanted brain machine interfaces, and mechanistic studies of brain function.

Computational Neuroscience

Computational Neuroscience seeks to understand and ultimately to emulate the computational principles and architectures of biological neural systems, perhaps the ultimate example of concomitant evolution of algorithms and hardware, thereby capturing their unique ability to sense, interpret, interact with and adapt to the external world, and their capability for fuzzy, non-Boolean computation. Contemporary neuroscience has identified many of the mechanisms that support neural information processing and cognition. Our present understanding of neural processing has reached a stage where fundamental advances can best be achieved by integrating existing ideas from neuroscience and computation into systems-level processing models. Indeed, given the scale and complexity of brain systems, computational simulation will be essential to identify the elements and explore the mechanisms of neural information processing. Ongoing work will result in computational systems that implement the computational principles and architectures of biological neural systems at appropriate scale and in real time, to enable efficient study of the emergent properties of neural systems. Detailed network simulations based on physiological data was identified as an exascale computing challenge in the DOE workshop on Extreme Biology.

Milestones FY14

- Develop linked computational models of retina and cortical visual areas and demonstrate role of neural dynamics in neural encoding, processing, and cognitive function
- Develop strategies for high performance FPGA/GPU implementations of cortical modules

Milestones FY15

- Develop and evaluate spiking models of early visual cortical areas including the primary visual area (V1) and the principal motion processing area (MT)
- Demonstrate the use of spike-based systems to detect correlation, infer causality, and implement probabilistic computation

Milestones FY16 and Beyond

- Implement neural network model systems to extract information encoded in synchronous oscillations of large neuronal populations
- Develop models of cortical motion processing for visual control of autonomous vehicles
- Develop general strategies for other modalities of neural pattern recognition, such as in natural language processing

Stretch Goals

- Develop model neural systems that can compete with human observers in perceptual processing tasks

Neuromorphic Computing

In order to achieve the computational performance, device densities, and power efficiency of biological brains on real-world problems, we are increasingly interested in neuromimetic systems that emulate biological neural function in specialized neuromorphic hardware or in distributed, heterogeneous hardware platforms. One objective is to build next generation computational systems to achieve intelligent processing of real-time, real-world data in portable, low-power devices. Another is to build distributed networks of sensors for ubiquitous surveillance of engineered structures or for applications such as border or perimeter monitoring. Our approach is to emulate the computational principles and architectures of biological neural networks in order to mimic their ability to reliably sense, monitor, interpret, navigate, interact with, and adapt to the external world, using inexpensive, low power fault tolerant systems. Synthetic neuromorphic “brains” that approach the performance of biological systems are the target of several DARPA/DoD/Intelligence community (IC)-sponsored programs. LANL is positioned to make an important impact in this arena both in traditional machine learning strategies as well as approaches that are strongly rooted in neuroscience. Neuromimetic computation also offers important lessons for high performance (even exascale) hardware systems, including:

- *Fault tolerance*–Neural information representation and processing are based on ensemble activity, so that single-point failures seldom matter
- *Reduced power*– "Spike"-based asynchronous communication and analog computation require lower power; but low-power digital computation may entail occasional errors and nondeterministic calculations
- *Self-Programming for massively parallel computation*–Neural systems learn through competition and generic learning rules with little explicit supervision, to achieve sparse, distributed, and statistically independent representations

Milestones FY14

- Implement a spiking model of retina on FPGA–demonstrate that the system exhibits dynamic encoding characteristic of biological retina
- Work with collaborators (U. Michigan, IBM, UCSD), to implement computation on spiking neuromorphic hardware and demonstrate the utility of such systems

Milestones FY15

- Develop and apply spiking models of cortical areas for FPGA- or GPU-based hardware
- Implement and demonstrate biologically inspired distribute networks for surveillance

Milestones FY16 and Beyond

- Implement linked spiking model of multiple cortical areas; explore the consequences of known feedback pathways for system dynamics and information processing activities

Stretch Goals

- Develop support for neuromimetic systems and applications for specialized HPC, and as a component of the exascale computing initiative within DOE

Application Areas

Theme: AstroInformatics

The way we gather information about the world was revolutionized when computers, a technology that had been around for more than 40 years, were linked together into a global network called the World Wide Web and real-time search engines like Google were created. The melding of modern information technology with autonomous robotic sensors is driving a similar revolution in observational science. Advanced information technology is also enabling new approaches for exploring the massive data volumes typically generated by modern simulations. It is clear that mastery of the new technology and its integration into practical systems will enable previously impossible discoveries in pure science and will be essential for constructing modern national security systems. The focus of the Astroinformatics Theme is the application of this modern approach both to fundamental scientific problems in astrophysics and cosmology and to applied problems in Space Situational Awareness (SSA).

Milestones and Goals for FY14

- Complete fabrication and successfully deploy an autonomous robotic instrument capable of making polarimetric measurements of Resident Space Objects and astrophysical transients; develop software for interpreting measurements.
- Construct a LANL catalog of photometric observations of Resident Space Objects
- Develop and deploy a next generation transient classification engine for the Palomar Transient Factory (PTF) collaboration
- Develop and deploy spectral classification algorithms for the new SED (Spectral Energy Distribution) Machine under development at Caltech for the iPFT collaboration
- Consolidate and expand the LANL role on the Large Synoptic Survey Telescope (LSST)—increase involvement in the astrophysical transients group; develop a plan for deploying a full sky monitor at the LSST site on Cerro Pachon in Chile
- Develop the initial designs and construct prototypes for the first instruments capable of persistent full sky monitoring from moving platforms
- Write scientific papers and white papers for potential external sponsors that demonstrate the power of autonomous telescope ecosystems for persistent surveillance

External Partners

- iPTF Collaboration
- LSST Collaboration
- NASA (Goddard, JPL, Marshall, Ames)
- Air Force Research Laboratory (AFRL)
- SNL/LLNL/ LBNL
- University of California (UCB)
- Caltech
- Las Cumbres Observatories Global Telescope Network (LCOGT)

Theme: Bioinformatics

The primary goals of bioscience at LANL are: 1) reducing the global biothreat, including both naturally occurring and artificial threats, and 2) understanding biological processes quantitatively to fulfill important national needs. These needs include understanding the interrelationships of climate and biological systems, developing new energy sources, and engineering novel, economically significant materials, based on biological principles.

Objective 1

Determine how to cope with the massive data streams (sequence, metagenomic, biosurveillance) that are being generated and require analysis. Global biosurveillance, metagenomics, biodefense, and public health all require data that is being produced in massive quantities, but which cannot be effectively exploited using current approaches.

Milestones FY14

- During FY 12/13 multidisciplinary teams, (Bioscience, IS&T), were established to address data acquisition, management, assimilation, and analysis, and apply novel algorithmic approaches including uncertainty quantification, machine learning, and visualization to the problems above. Considerable internal "buy-in" was obtained for addressing the problem of "multi-omics" data integration. A number of internal and external proposals have been prepared and submitted in this area. These teams continue to work together to seek additional funding, both internal and external. A focused effort in global biosurveillance was established which should continue to provide guidance for setting IST goals.
- Leverage successes in global biosurveillance to generate opportunities for additional funding (Defense Threat Reduction Agency, DTRA, DHS)

Milestone FY15 and beyond

- During FY12/13 inroads were made in establishing connections between computational biology and the needs of the intelligence community. We need to continue to obtain a better understanding of the needs of the intelligence community and develop strategies to support these needs and obtain additional funding.

Stretch goals

- Develop principled methods for integrating data
- Make advances in natural language processing by partnering with external entities

Objective 2

Expand our quantitative understanding of biological systems to optimize existing biological functions or develop new ones to enhance our ability to tackle large-scale problems by integrating diverse capabilities that are already present at LANL.

Milestones FY14

- During FY12/13 we began to develop large-scale integrated quantitative pathway models for hosts, pathogens and host-pathogen interactions. Proposals in this area were developed and submitted both internally and externally.
- We need to expand currently funded efforts in biological network modeling and diversify this funding to include other investigators and other sponsors.

- Submit experimental/computational proposals that incorporate unique Laboratory experimental capabilities
- Seek a collection of proposals aimed at the design of a regulator element such as a transcription factor or surface-displayed binding element, assemble multidisciplinary teams (Bioscience, IS&T) to carry out this work based the work on computationally driven experimental methods

Milestones FY15 and Beyond

- Demonstrate a clear adaptive design cycle for protein biodetection systems. The goal is to establish LANL as a world leader in biodetection by design.
- Make advances in simulating enzyme activity. Experimental capabilities at the Laboratory in deep sequencing and high throughput assay systems (e.g. cellulose activity screens at genomic scale) are generating data which can be correlated with individual amino acid perturbations. Establishing an effort in this area will assist in integrating Laboratory experimental and computational communities.
- A potential new area is using biopolymers to construct advanced materials. For example, the Baker Lab has made advances that will help enable design of self-assembling DNA frameworks with the functional properties of proteins. Materials scientists can now functionalize protein block polymers and surface protein assemblies. Modeling and simulation activities are currently following and not leading these efforts.

Stretch Goal

- Large scale simulation of self-assembling peptide frameworks

Theme: Climate Informatics

Climate informatics is the modeling, prediction, and analysis of climate and Earth systems, and the support and processing of observations and observation datasets. LANL is unique in its development and application of high-performance, high-resolution models of oceans, sea ice, and land ice, with close ties to the climate modeling community (especially the Community Earth System Model [CESM]). LANL has a significant capability in modeling climate system impacts to environment, populations, and infrastructure (including coastal impacts) and the capability to translate this into decision support. LANL deploys novel *in situ* and remote sensing resources to monitor key climate metrics, and combines these with community observations to develop an understanding of the environment and climate systems. LANL has a particular focus on issues surrounding global and regional sea-level rise, high latitudes and polar climate, and southwestern climate. This Climate Informatics theme is in alignment with and supportive of LANL's Energy Climate Interactions (ECI) Initiative.

Milestones FY14

- Maintain and review Energy Climate Impacts (ECI), emphasizing IS&T strategy, in order to define current issues and key projects for LANL capabilities and programs in the interlocking areas of energy security, climate, and impacts efforts across key customers at LANL
- Make large-scale analytics strategies available to be used for model validation and evaluation, and scientific application of model detailed output and other Earth system spatio-temporal datasets (larger IS&T goal: analysis of massive scientific datasets)

Milestones FY15

- MPAS-O next-generation multi-scale global ocean model available for model-based predictions and support of scientific (Climate, Ocean, and Sea Ice Modeling, COSIM)
- Workshop bringing together LANL expertise and ongoing IS&T work on infrastructure impacts of climate change, including informatics and analytics strategies and requirements
- Demonstration of combining uncertainty across multiple models, using fast reduced-complexity models (Office of Science, *Combining System and Model Dynamics to Learn About Climate Uncertainties [Urban]*)

Milestones FY16 and Beyond

- Uncertainty analysis and decision models connected to policy areas in response to climate change, impacts, and energy infrastructure issues (programs in development) targeting a strategic hire interfacing modeling and simulation with infrastructure and impacts relevant to climate change
- Enhanced LANL staff participation in climate modeling, climate change, and impacts and adaptation decision support working groups at a planning and advisory level

Stretch Goals

- Global multi-scale coupled climate models available for investigation of local to regional impacts analysis, attribution of anthropogenic change and natural variability; running at scale on exascale hybrid architectures
- Recognized expertise (go-to) in model-based impacts assessment and decision support for planning related to natural resources, societal resilience, and infrastructure adaptation in response to climate change and weather extremes, especially in the area of detailed models of critical infrastructure, energy security, and water impacts related to national and global security
- Develop framework and principles for monitoring of simulation model enhancement with limited and uncertain observations, including strategies for future climate inference and uncertainty quantification (UQ), and experimental design
- Leadership in application of analysis of massive space-time datasets using exascale high performance computers and data-intensive computing resources, including storage resources that can make these datasets available for fast and flexible access

External Partners

- Program maintains strong connections with various Department of Energy (DOE) and academic faculty–interactions with a large diverse community is a necessity
- National Center for Atmospheric Research (NCAR)
- University of Texas Predictive Science Academic Alliance Program (PSAAP)/Center for Predictive Engineering and Computational Sciences (PECOS) and affiliated staff
- DOE labs, especially Lawrence Livermore National Laboratory (LLNL), Oak Ridge National Laboratory (ORNL), Lawrence Berkeley National Laboratory (LBNL), and Pacific Northwest National Laboratory (PNNL)
- University of California Santa Cruz (UCSC)

- Potential partner to develop: NCAR/NSF, NOAA, Navy
- Cultivate collaborations and partnership with climate policy institutions and community

Existing Efforts

- LANL COSIM team funded through DOE/Office of Biological and Environmental Research (BER), DOE/Scientific Discovery through Advanced Computing (SciDAC), and others
- DOE/BER Predicting Ice Sheet and Climate Evolution at Extreme Scales (PISCEES)
- DOE/SciDAC QUEST Institute
- LDRD DR and Exploratory Research (ER) projects

Theme: High Performance Computing Operational Informatics

Large HPC data centers are tremendous generators of data. In addition to the data produced by applications, the supercomputers themselves and the facilities operating them can produce a tremendous number of discrete data points. At LANL, approximately 180,000 compute cores comprising 15 supercomputers and their associated software subsystems and physical environments produce data that must be collected, organized, and analyzed. Individual data streams provide basic information about system or subsystem health and performance, but the analysis of interacting data streams is necessary to generate a more complete understanding of the interaction of complex systems.

Utilization Metric Variations

Milestone FY14

- Cross-reference source data for more complete system information

Milestone FY15

- Extend efforts from correctness to utilization data sources

Power Usage Investigations

Milestones FY14

- Publish more detailed power fluctuation projections for Trinity
- Gather requirements for power-aware job scheduling

Milestone FY15 and Beyond

- Continue projects with more finely-grained power-usage data sources

System/User Data Integration

Milestone FY14

- Design integration of select system and application performance data

Milestone FY15 and Beyond

- Integrate system and application performance data

Theme: Materials Informatics

Materials Informatics (MI) can be broadly defined as the application of computational methodologies to processing and interpreting materials science and engineering data. Specialized informatics tools for data capture, management, analysis, and dissemination, together with advances in computing power, coupled with computational modeling and simulations, can potentially enable accelerated materials development and discovery that can save millions of dollars. More broadly, MI can be interpreted as a learning and co-design process that, when properly optimized, crosses materials and information science boundaries to uncover fundamental knowledge that is the basis of physical, mechanical, and engineering properties.

Materials Informatics can be divided into two main parts: data management and knowledge discovery. New experimental methods, such as combinatorial materials science, are generating large amounts of structural and property data. Computational methods are being developed to classify and calculate the structure and properties of materials. Thus, experiment and theory generate large amounts of data that are highly heterogeneous and fill gaps in materials property space and provide routes to estimating properties and improve property prediction or material selection. For a successful design and applications strategy, the challenge is to optimize data and knowledge for desired results. With the cost of computing shrinking rapidly, the accuracy of computations and results from data-mining algorithms are improving. Computational codes, be they based on ab initio, molecular dynamics (MD), or finite element methods, are enhancing our ability to tackle ever-larger systems and address complex materials issues by generating extensive theoretical databases. Materials Informatics should allow us to answer questions of the following type:

- Given a database of known materials, what unknown material has certain desirable properties most similar to a known one?
- Can materials be clustered or partitioned into classes so as to discover useful patterns?

While informatics has impacted other scientific areas, it remains at a rudimentary stage with relatively little penetration in materials science. Therefore, the time is ripe to explore the available diverse techniques in informatics to design and discover new materials.

Data Analysis, Management

Milestones FY 14 and Beyond

- Second conference in Mapping the Materials Genome series entitled “Materials Informatics: Information Science for Materials Design” held at LANL – Februrary 2014
- Initiate codesign loop proposed in LDRD-DR by establishing pertinent data bases and using state-of-the-art statistical tools for classification and regression

Required Resources

- The funded LDRD-DR will focus on the development of analysis tools on specific data sets. The wide applicability of these tools to materials and other data across the LANL campus suggests that there should be a centralized database that links with data resources and domain knowledge across LANL.
- Help with recruitment of students, post-docs and entry level staff that can combine expertise in machine learning, applied mathematics with domain knowledge in this burgeoning field
- Continued annual workshop support

External Partners

- Institute for Combinatorics Discovery, Iowa State University
- Genomic Signal Processing Laboratory, Texas A&M University

Knowledge Discovery

Milestones FY 14 and Beyond

- Conference: Nanoinformatics and the Materials Genome–(in collaboration with the National Nanotechnology Manufacturing Initiative)
- Conference: Materials Discovery with Alternative Materials

Stretch Goal

- Develop the Materials Data Foundry in conjunction with the FY14 funded LDRD-DR

Required Resources

- DOE Basic Energy Sciences (BES) funding
- Support for students, post-docs and entry-level staff
- Help in generating greater interest and awareness, as well as sources of materials centric data across LANL

External Partners

- Iowa State University
- Massachusetts Institute of Technology (MIT)
- Stanford
- Ames Lab
- Boeing
- Carnegie Mellon University (CMU)
- Texas A&M University

- Iowa State University
- MIT
- CMU
- Texas A&M University

Theme: Smart Grid

Increased deployment of new technologies, e.g., renewable generation and electric vehicles, is rapidly transforming electrical power networks by crossing previously distinct spatiotemporal scales and invalidating many traditional approaches for designing, analyzing, and operating power grids. This trend is expected to accelerate over the coming years, bringing the disruptive challenge of complexity, but also opportunities to deliver unprecedented efficiency and reliability.

LANL's vision of research and development in this area is centered on three applied mathematics pillars: complex systems theory, control theory, and optimization theory (Fig. 3). These pillars are inherently interconnected by emerging problems in power grids. Existing approaches that ignore one or more of these pillars are increasingly incomplete because they disregard fundamental emergent couplings. Instead, LANL's unique approach integrates these pillars through an iterative, multifaceted approach. A full integration of these three pillars produces the necessary mathematical tools to achieve many far-reaching goals, including, but not limited to, a fully automated, real-time monitoring, analysis, and control system for large-scale electric power grids.

Complex Systems—LANL's approach begins with complex systems theory, where the distinct scales of the complete network are specified and the basic static, dynamic, and stochastic phenomena are analyzed. We iteratively partition, or *"deconstruct"* the network and its processes into non-separable spatiotemporal scales and identify separate sub-problems and the crucial couplings between them. Several methods of model reduction are applied at this stage, e.g., smoothing over spatially discrete network flows/injections by conversion of network flow models to Ordinary (ODE) or Partial Differential Equations (PDE), construction of hybrid dynamical system representations, and rare event analysis to identify a small number of the most probable, yet damaging, network fluctuation modes out of a continuous space of possibilities. Models underlying these methods and phenomena determine scale-specific optimization and control sub-tasks and formulate important practical engineering problems in a mathematically sound form. These well-formulated control problems are passed to the control theory pillar.

Control Theory—The structure of the complex network formulations is influenced by state- of-the art control concepts. The mathematical building blocks of control theory must also develop to rigorously handle new and difficult problems, e.g., large-scale, distributed control problems with hierarchical constraints that require actions over a wide range of time scales with varying degrees of information available. Problems that are particularly large, discrete, nonlinear, and/or constrained are passed to optimization theory.

Optimization Theory—Although related to control theory, optimization theory is better suited to address complicated nonlinear, discrete problems. Their size and structure requires advances in the mathematical foundations of optimization theory and algorithms to improve computational efficiency and accuracy, e.g., novel relaxation methods and approximation techniques that produce effective bounds in practicable CPU time and new methods for multi-level, robust, and stochastic optimization with chance constraints.

The three pillars are then implemented in a hierarchical and iterative manner to achieve a solution that is greater than each pillar on its own.

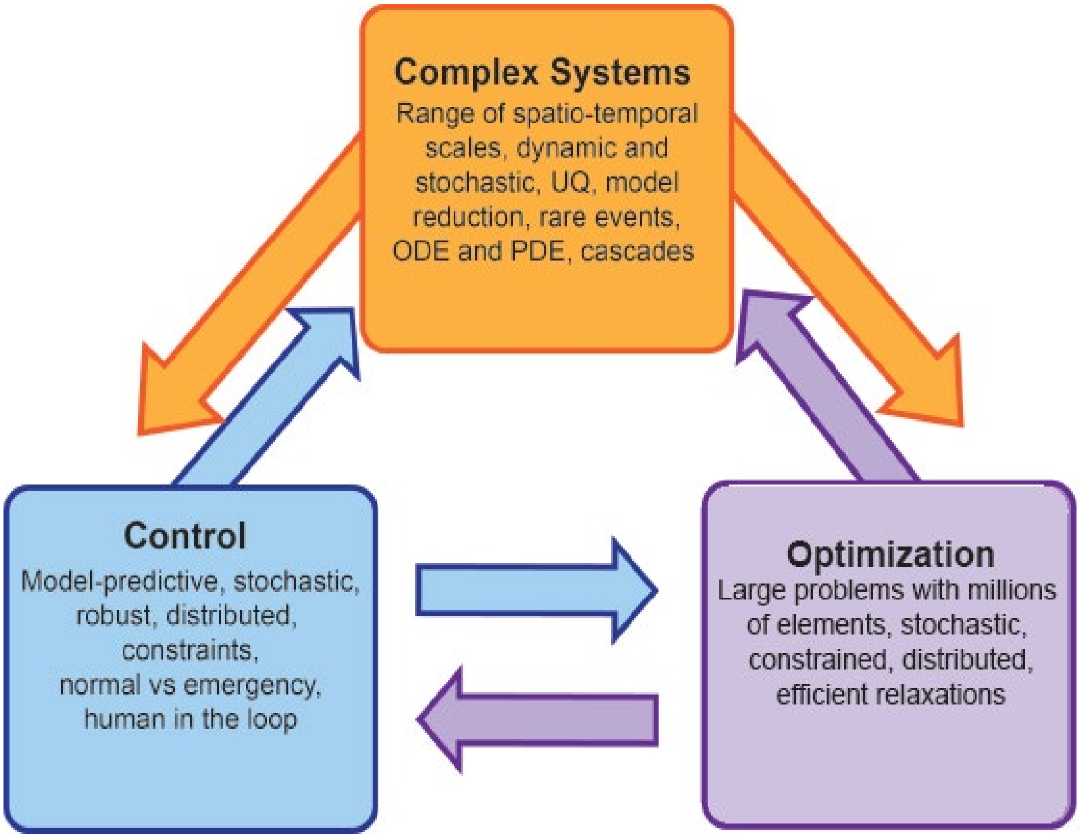


Figure 3. The three applied math pillars of the AMPS Center and the iterative flow of model development and solution.

Goals for FY14 and Beyond

- Establish a LANL-led program in "Grid Resilience by Design" ($1-10M depending where/how the opportunity shows up, leveraging both Grid Science and NISAC programs)
- Establish collaboration with PNNL on "Next Generation Energy Management System" (~$1M again, depending on where/how the funding shows up, leveraging Grid Science algorithms)
- Establish collaboration with NREL in "Energy Systems Integration" (NREL lead, ~$1M to LANL again depending on where/how the funding shows up, leveraging NISAC modeling and software architecture and integration)
- Expand DTRA projects to include 6.2 funding to support DTRA's current 6.1 funding at LANL

- Establish collaborative projects with industry on using game theory to design cyber-physical control systems

Required Resources

- Institutional backing/support of Goals #1-3 in planning conversations with other national labs
- Cross SPO-AE/GS program office support for Goals 1, 4, and 5
- Program Development funds to support the travel necessary for Goals 1-5

Theme: Structural Health Informatics

The process of implementing a damage-identification strategy for aerospace, civil, and mechanical engineering infrastructure is referred to as structural health monitoring (SHM). This process involves the observation of a structure or mechanical system over time using periodically spaced dynamics response measurements, the extraction of damage-sensitive features from these measurements, and the statistical analysis of these features to determine the current state of system health. For long-term SHM, the output of this process is periodically updated information regarding the ability of the structure to continue to perform its intended function in light of the inevitable aging and damage accumulation resulting from the operational environments. Under an extreme event, such as an earthquake or unanticipated blast loading, SHM is used for rapid condition screening. This screening is intended to provide, in near real time, reliable information about system performance during such extreme events and the subsequent integrity of the system.

At LANL, hardware and software solutions to SHM problems are based on a four-step statistical pattern recognition paradigm. This four-step process comprises of: 1) operational evaluation; 2) data acquisition, normalization, and cleansing; 3) feature selection and information condensation; and 4) statistical model development for feature classification. Research activities include wireless sensor nodes, energy harvesting for embedded sensing, sensor diagnostics, sensor system optimization, robotic sensor systems, model-based and data-based approaches to damage-sensitive feature extraction, machine learning approaches to data normalization and feature classification, information-gap assessment of SHM system robustness, and the SHMTools software package. Applications include highway bridges, space-situational awareness telescopes, aerospace structures, amusement park rides, navy ships, wind turbines, pipelines, and rotating machinery.

Milestones FY14 and Beyond

- Workshop: Integrating SHM Technology into a Cradle-to-Grave System State Awareness Capability
- Proposal Development: integrate SHM with control systems, damage evolution predictions, and reliability-based decision making techniques
- Expand applications to third-party engineered systems (DoD, DOT, Private Industry)
- Continue to focus at least one Los Alamos Dynamics Summer School project on SHM
- Continue to identify and pursue high-quality funding opportunities
- Continue to host visits by international scholars from partner universities

Stretch Goal

- Integrate SHM technology into a flagship System State Awareness Engineering Center comparable to Sandia's MESA and Robotics facilities

Required Resources

- Funded projects that motivate the requisite multi-disciplinary technology development
- Line and program management internal and external promotion of this technology
- Recruitment of entry-level staff that have the educational background to work on multi-disciplinary technology development
- Laboratories to support the requisite multi-disciplinary SHM technology development

External Partners

- University of California San Diego (UCSD)
- Chungbuk National University, Jeonju, South Korea
- AGH University of Science and Technology, Krakow, Poland
- University of Sheffield, UK
- University of Bristol, UK

Existing Efforts

- Intelligent Wind Turbine LDRD-DR
- SHMTools software development
- Raptor telescope SHM feasibility study

LANL Organizations Supporting the IS&T Pillar

IS&T Institute (ISTI)

Since 2008 The IS&T Center has been coordinating the Integrating IS&T for Prediction Science Pillar. In the summer of 2012 the IS&T Center was incorporated into the IS&T Institute (ISTI). The mission of the new ISTI is to enhance LANL's capabilities in IS&T by:

- Fostering science, technology, and engineering to address LANL, DOE, and national IS&T needs
- Coordinating Laboratory-wide IS&T efforts and aligning them with future needs
- Positioning LANL to compete in a larger market for funding in this important and growing area

The ISTI provides a connection to program management for capability needs and provides IS&T integration and support for mission-critical centers and activities such as MaRIE. The Institute helps to coalesce and strengthen nascent IS&T capabilities, as well as leverage existing, mature capabilities. It takes part in communicating IS&T activities internally and externally and advises Laboratory management on IS&T investment strategies. In addition to exploiting the synergy between CS/Math for HPC and CS/Math for IS&T, the ISTI raises visibility of IS&T as a technical discipline at LANL.

The ISTI works with LANL program management to create opportunities for IS&T funding and with line management and technical staff to facilitate the execution of IS&T projects. The Institute directly funds some program development activities in IS&T and works with line management to recommend areas for strategic LDRD funding.

Collaborative Research

ISTI engages in, implements, and/or broadly fosters collaborative research, capability and workforce development, technical exchange, program development activities, and institutionally-focused strategic planning. ISTI consists of a collection of highly successful, strategic partnerships with leading research universities focused on IS&T topic areas of critical importance to LANL. IS&T, including but not limited to data management at scale, high performance computing, data-intensive computing, computational co-design, reliability and resilience at scale, algorithms and methods, informatics, and multi-core and hybrid computing, is a key capability that underlies overall scientific leadership and achievement across scientific domain areas. Through its strategic partnerships, ISTI works to facilitate and enhance scientific collaboration and scholarship, and workforce development (recruitment, retention, and revitalization) through a breadth of programs.

Research Library

The Research Library (RL) has recently become involved with the NSF DataONE effort: http://www.dataone.org/. This effort explores both socio-cultural issues and cyber infrastructure issues around sharing data. RL anticipates participating in many aspects, helping to link DOE and other relevant resources, and potentially joining as a "member node" that would include sharing data from LANL. Becoming a member node would require sharing some kind of open environmental data with the DataONE network. In the realm of data "management," RL has established a data team and is getting trained (via DataONE resources, etc.) about what this means—first is education of RL personnel, then a rollout of an education program for LANL researchers.

Center for Nonlinear Studies

CNLS provides an interdisciplinary environment where postdocs, students, and visitors interact with LANL scientists to stimulate and advance new research directions. Information science is a major focus of the research at CNLS (http://cnls.lanl.gov/). CNLS will continue to provide space for the IS&T Seminar Series and will help with and contribute to ISTI workshops.

Advanced Computing Solutions

The Advanced Computing Solutions (ACS) Program's Office focus is to enable LANL to become nationally recognized as the capability leader in information security and networking science, by predicting and solving critical problems in the cyber domain using novel and practical solutions. Information security and networking science are intended to be growth elements of the Laboratory with diverse customers leveraging excellence in the Laboratory's own cybersecurity. ACS's mission and goals are to:

- Partner with other LANL capabilities to leverage and successfully advance the Laboratory and its programs
- Support LANL management in identifying and solving the Laboratory's information security challenges

- Prove concepts tangibly and recommend practical solutions to existing and potential customers
- Comprehensively understand and articulate issues and second-order effects in both the problem and solution space
- Formulate and appropriately translate relevant cyber research and engineering problems into opportunities for other LANL capabilities
- Effectively match leading-edge technology innovations and equally emergent information-based threats
- Create and nurture an agile organization with sufficient depth and breadth for sustained growth

ASC Program Office

The ASC program was established in 1995 to help shift from test-based confidence to simulation-based confidence. Its mission is to analyze and predict the performance, safety, and reliability of nuclear weapons and certify their functionality. The tri-lab collaborators also work in partnership with computer manufacturers and in alliance with several of the nation's major universities. The ASC program office will continue to ensure that the Integrating IS&T for Prediction pillar impacts the NW program.

LANL Office of Science Program Office

LANL's Office of Science programs interface with the Laboratory's IS&T Strategic and Implementation Plans in at least three ways (mission foci aligned with the DOE Offices of BES, BER, and science foci aligned with ASCR). The IS&T mission foci on predicting materials behavior and on energy/climate interactions align strongly with our portfolios in the DOE SC, BES and BER, respectively. Co-design, broadly defined to include the coupling of experiment to theory and IS&T, is key to our efforts, as evidenced by Exascale Co-design Center proposals led by LANL in each of these areas. Further, LANL's ASCR portfolio, which has significant activities in the science of co-design and of UQ aligns well with the IS&T science foci.

The G.T. Seaborg Institute for Actinide Science

Understanding the full equation of state for such materials of national interest as plutonium (Pu) is beyond the current capabilities of experiment. In areas where we cannot measure, we must model and simulate. The complexity of the electronic structure of Pu will stress the most advanced computational resources as models and theory move to incorporate thermal vibrations and alloy elements in supercells; these require both huge computing resources and advanced algorithm approaches. The Seaborg Institute supports postdoctoral research fellows in these areas, where a strong overlap with the IS&T implementation plan exists.

The Engineering Institute

The Engineering Institute (EI) is an education and research collaboration between LANL and the UCSD Jacobs School of Engineering. The technical thrust of the EI is the promotion of emerging multidisciplinary engineering research that integrates predictive modeling, advanced sensing capabilities, and information technology. Originally these technologies were integrated to address the general field of “damage prognosis,” which is concerned with assessing the current condition and predicting the remaining life of aerospace, civil, and mechanical engineering systems. Recently the EI’s research portfolio has expanded to include broader engineering research activities that can be more generally described as development of cyber-physical systems. Numerous graduate level classes offered through the EI and much of its research portfolio is focused on IS&T. The EI also has developed close collaborations with UCSD’s California Institute for Information Technology and Telecommunications.

LDRD Office

The Information Science and Technology Grand Challenge (one of five) addresses the development of breakthrough mathematics, computer science, and underlying technologies required to extract information, knowledge, and insight from data. This Focus Area is closely aligned with the Integrating IS&T for Prediction Pillar and closely supports the Materials for the Future and Science of Signatures Pillars. IS&T will continue to be an emphasis in both the DR and ER categories.

Institutional Computing

The Institutional Computing Program was developed to provide production computing resources for open and collaborative science at LANL. Institutional Computing provides an equal opportunity for access to these resources to every scientist and engineer at LANL through a competitive, peer-reviewed proposal process.

LANL encourages the creative use of computer resources to enhance our research and development activities. To promote such broad, open endeavors, the Institutional Computing project currently provides almost 40 million CPU hours per year without cost to the user.

(See the Institutional Computing Five-year Plan.)

Glossary

Term	Meaning
ACS	Advanced Computing Solutions
ActivitySim	Agent-based model using synthetic US population
ADE	Associate Directorate for Engineering and Engineering Sciences
ADIT	Associate Directorate for Information Technology
ADTSC	Associate Directorate for Theory, Simulation, and Computing
AFSOR	Air Force Office of Scientific Research
ANL	Argonne National Laboratory
ASC	Advanced Simulation and Computing
ASCEM	Advanced Simulation Capability for Environmental Management
ASCR	Advanced Scientific Computing Research
ASET	Architectural Simulation and Exploration Tool
AWE	Atomic Weapons Establishment
BER	Office of Biological and Environmental Research
BES	Office of Basic Energy Sciences
BRAIN	Brain Research Through Advancing Innovative Neurotechnologies
CASL	Consortium for Advanced Simulation of Light water reactors
CCS	Computer, Computational, and Statistical Sciences
CCSI	Carbon Capture Simulation Initiative
CEA	French Alternative Energies and Atomic Energy Commission
CESM	Community Earth System Model
CIO	Chief Information Officer
CMU	Carnegie Mellon University
CNLS	Center for Nonlinear Studies
COSIM	Climate, Ocean, and Sea Ice Modeling
CPAM	Computational Physics and Applied Math
CRADA	Cooperative Research and Development Agreement
CyberSim	Simulations of software vulnerabilities and the resulting malware propagation in online social networks
DARPA	Defense Advanced Research Projects Agency
DataONE	Data Observation Network for Earth
D2D	Data to Decisions
DDES	Distributed discrete even simulation
DHS	Department of Homeland Security
DoD	Department of Defense
DOE	Department of Energy
DTRA	Defense Threat Reduction Agency

Term	Meaning
EC EUDAT	European Commission European Data Infrastructure
ECI	Energy climate impacts
ECL	Enterprise Control Language
EES	Earth and Environmental Science Division
EI	Engineering Institute
fMRI	Functional magnetic resonance imaging
FPGA	Field-programmable gate array
FRE	Functional Reasoning Engine
FSU	Florida State University
GHG	Greenhouse gas
GHGIS	Greenhouse Gas Information System
GIAC	Graphical Interface Aggregate Control
GPGPU	General-purpose graphics processing units
GPU	Graphics processing unit
GS	Global Security
HACC	Hardware Accelerated Cosmology Code
HPC	High Performance Computing
IARPA	Intelligence Advanced Research Projects Activity
IC	Institutional Computing
iCORDI	International Collaboration on Research Data Infrastructure
iPFT	Intermediate Palomar Transient Factory
ISTI	Information Science and Technology Institute
IT	Information Technology
LANL	Los Alamos National Laboratory
LBNL	Lawrence Berkeley National Laboratory
LCOGT	Las Cumbres Observatory Global Telescope
LDRD	Laboratory Directed Research and Development
LLNL	Lawrence Livermore National Laboratory
LSST	Large Synoptic Survey Telescope
MaRIE	Matter-Radiation Interactions in Extremes
MD	Molecular dynamics
MGH	Massachusetts General Hospital
MI	Materials informatics
MIITS	Multi-scale Integrated Information and Telecommunications System
MIT	Massachusetts Institute of Technology
MPAS-O	Next-generation ocean model
MT	Middle temporal part of the brain
NCAR	National Center for Atmospheric Research

Term	Meaning
NIH	National Institutes of Health
NISAC	National Infrastructure Simulation and Analysis Center
NIST	National Institute of Standards and Technology
NNSA	National Nuclear Security Administration
NSF	National Science Foundation
NW	Nuclear weapons
OCIO	Office of the Chief Information Officer
ODE	Ordinary Differential Equations
ORNL	Oak Ridge National Laboratory
OSTP	Office of Science and Technology Policy
PADGS	Principal Associate Directorate for Global Security
PADSTE	Principal Associate Directorate for Science, Technology and Engineering
PDE	Partial Differential Equation
PECOS	Predictive Engineering and Computational Sciences
petaFLOP	Measure of a computer’s processing speed; expressed as a thousand trillion floating point operations per second
PI	Principal Investigator
PISCEES	Predicting Ice Sheet and Climate Evolution at Extreme Scales
PSAAP	Predictive Science Academic Alliance Program
PSP	Predictive Science Panel
PTF	Palomar Transients Factory
Pu	Plutonium
PV	Photovoltaic
QIS	Quantum information systems
QIS	Quantum Information Science
QUEST	Quantification of Uncertainity in Extreme Scale
R&D	Research and Development
RD100	R&D Magazine R&D 100 Awards
RL	Research Library
SciDAC	Scientific Discovery through Advanced Computing
SDAV	Scientific Data Management, Analysis, and Visualization
SED	Spectural Energy Distribution
SNL	Sandia National Laboratories
SOO	Strategic Outcomes Office
SPO-AE	Science Programs Office – Applied Energy
SQUID	Superconducting Quantum Interference Devices
SQuINT	Southwest Quantum Information and Technology
SSA	Space situational awareness

Term	Meaning
TRANSIMS	Transportation Analysis and Simulation System
UCD	University of California Davis
UCR	University of California Riverside
UCSC	University of California Santa Cruz
UCSD	University of California San Diego
UI	User Interface
UK	United Kingdom
UNC	University of North Carolina
UNM	University of New Mexico
UOH TAMU	University of Texas- Texas A&M Universtiy
UPSIDE	Unconventional Processing of Signals for Intelligent Data
UQ	Uncertainty quantification
UT	University of Tennessee
V1	Primary visual area
WFO	Weapons Facility Operations
WMD	Weapons of mass destruction
XCP	Computational Physics
XTD	Weapons Design